MATERIAL DIDÁCTICO
MATEMÁTICAS

10

ESTRATEGIAS PARA MEJORAR LA APLICABILIDAD DE MÉTODOS ITERATIVOS QUE UTILIZAN DIFERENCIAS DIVIDIDAS

José Antonio Ezquerro Fernández
Miguel Ángel Hernández Verón





ESTRATEGIAS PARA MEJORAR LA APLICABILIDAD DE MÉTODOS ITERATIVOS QUE UTILIZAN DIFERENCIAS DIVIDIDAS

MATERIAL DIDÁCTICO

Matemáticas

nº 10

José Antonio Ezquerro Fernández

Miguel Ángel Hernández Verón

ESTRATEGIAS PARA MEJORAR LA APLICABILIDAD DE MÉTODOS ITERATIVOS QUE UTILIZAN DIFERENCIAS DIVIDIDAS

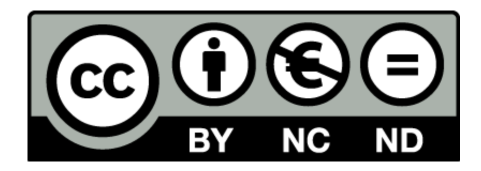
UNIVERSIDAD DE LA RIOJA

SERVICIO DE PUBLICACIONES

2014

EZQUERRO FERNÁNDEZ, José Antonio
Estrategias para mejorar la aplicabilidad de métodos iterativos que utilizan diferencias divididas [Recurso electrónico] / José Antonio Ezquerro Fernández, Miguel Ángel Hernández Verón. – Logroño : Universidad de La Rioja, Servicio de Publicaciones, 2014.
VII, 145 p. – (Material didáctico. Matemáticas ; 10)
ISBN 978-84-695-9421-6
1. Métodos iterativos (Matemáticas). 2. Ecuaciones diferenciales no lineales. 3. Resolución numérica de ecuaciones. I. Hernández Verón, Miguel Ángel. II. Título. III. Universidad de La Rioja. Servicio de Publicaciones. IV. Serie.

519.6
517.9
PBKS – IBIC 1.1



-

publicaciones.unirioja.es

E-mail: publicaciones@unirioja.es

ISBN 978-84-695-9421-6

Edita: Universidad de La Rioja, Servicio de Publicaciones

A María, la luz.

José A.

A Mercedes, por la vida tan intensa
que me ha proporcionado durante
estos últimos más de treinta años.
Por la ternura y el cariño
que me da cada día.

Michel

Prólogo

Uno de los problemas más antiguos de las matemáticas, y por extensión de las ciencias y las ingenierías, es la resolución de ecuaciones no lineales; y, en particular, la aproximación de las soluciones de las ecuaciones con suficiente exactitud. Así, nos podemos encontrar con una gran cantidad de problemas que se pueden reescribir en términos de ecuaciones a resolver, lo que conduce a una gran variedad de ecuaciones: sistemas finitos e infinitos, ecuaciones diferenciales ordinarias y en derivadas parciales sujetas a condiciones iniciales o de contorno, o una combinación de ambas, y ecuaciones integrales o integrodiferenciables. Además, generalmente, estas ecuaciones son no lineales. Las variables de las ecuaciones pueden ser números reales o complejos (ecuaciones algebraicas simples con una sola variable), vectores (sistemas de ecuaciones) o funciones (ecuaciones diferenciales o integrales).

Sin embargo, desde el punto de vista del análisis funcional, todas estas ecuaciones se pueden expresar en función de operadores que aplican algún espacio lineal en otro espacio lineal, de manera que las soluciones buscadas son elementos o puntos del espacio correspondiente. En consecuencia, los esquemas numéricos que, en este marco tan general, aproximan las soluciones de esta gran variedad de problemas conducen al desarrollo de métodos efectivos y fiables que aproximan soluciones, con suficiente exactitud, de las ecuaciones en el espacio original o en un espacio relacionado. Excepto en casos especiales, los esquemas numéricos de resolución más comúnmente utilizados son métodos iterativos que proporcionan una sucesión que converge, a partir de una o varias aproximaciones iniciales, a una solución de la ecuación a resolver. Como estos métodos tienen la misma estructura recursiva, los hemos introducido y estudiado dentro de un marco general: el de los espacios de Banach.

Durante los últimos años, los estudios de los métodos iterativos incluyen un esfuerzo sustancial en la identificación de propiedades que garanticen su convergencia en algún sentido. Una gran cantidad de estos resultados consisten en aplicar recursivamente métodos iterativos punto a punto o con memoria. En este texto, analizamos algunos de estos métodos.

Este texto es fruto de la tarea docente de sus autores, a lo largo de los últimos años, en cursos de doctorado enfocados a la resolución de ecuaciones no lineales para titulados en Matemáticas e Ingeniería. Su contenido va dirigido a estudiantes titulados en carreras científicas y técnicas que deseen conocer, de manera asequible, sencillos métodos iterativos que les permitan afrontar, con sencillez y rigor, cuestiones relacionadas con la resolución de ecuaciones no lineales. Hemos procurado, en todo momento, huir de un texto reducido a una colección de métodos iterativos. Hemos insistido en desarrollar técnicas de demostración asociadas a la materia aquí presentada, así como en el planteamiento de los problemas, en los conceptos fundamentales que intervienen, en las estrategias de resolución y en las cuestiones relacionadas con la convergencia. Se ha buscado un equilibrio entre la claridad del desarrollo y la profundidad de los conceptos. Y uno de nuestros objetivos ha sido un acercamiento de las matemáticas puras y aplicadas, que, a menudo, son vistas por los estudiantes sin conexión.

El principal objetivo de este texto es presentar una visión general de algunos resultados teóricos básicos sobre ecuaciones no lineales en espacios de Banach, así como analizar algunos de los métodos iterativos que resuelven numéricamente estas ecuaciones. También esperamos proporcionar un texto para cursos superiores del análisis numérico relacionado con estos tópicos y, para facilitarlo, hemos tratado de hacer un texto tan autosuficiente como nos ha sido posible. Para ello, hemos escrito un primer capítulo introductorio y probado gran parte de los resultados con todo detalle.

Mientras que las clases se pueden desarrollar en un contexto más o menos riguroso, hemos escrito este texto utilizando métodos y hábitos de argumentación pedagógicos. Los prerrequisitos del texto son el cálculo, el álgebra lineal y nociones menores de programación computacional. También cierta familiaridad con las técnicas básicas para resolver ecuaciones simples de una variable, así como de sistemas de ecuaciones lineales, resulta beneficioso. Si que es cierto que los estudiantes con conocimientos de los principios del análisis numérico tienen cierta ventaja, ya que los esquemas y conceptos generales se siguen fácilmente si ya se conocen métodos y casos particulares. Sin embargo, dicho conocimiento no es prerrequisito indispensable para entender el material aquí tratado. Ocurre lo mismo con los conocimientos del análisis funcional y de los espacios de Banach en particular. Todo el material necesario y requerido para la teoría desarrollada a lo largo del texto está recogido en el capítulo 1.

Dos son principalmente las cuestiones que más nos han preocupado y a las que hemos tratado de dar respuesta. Por una parte, el análisis de la convergencia de métodos iterativos, con memoria y punto a punto, que utilizan diferencias divididas cuando se aplican a la resolución de ecuaciones no lineales. Para ello, presentamos tres técnicas diferentes de demostración: el principio de la mayorante de Kantorovich, una técnica basada en relaciones de recurrencia y una modificación novedosa de la técnica anterior. Por otra parte, la obtención de aproximaciones iniciales suficientemente buenas para que los métodos iterativos aquí considerados converjan empezando en ellas.

Hemos dividido el texto en tres partes con cinco capítulos y el texto trata de ser autosuficiente. Cada capítulo contiene varios resultados que se han ilustrado con aplicaciones a la resolución de sistemas de ecuaciones no lineales que surgen de procesos de discretización de ecuaciones diferenciales e integrales no lineales que aparecen con cierta frecuencia en las ciencias y las ingenierías.

La primera parte del texto consta de un solo capítulo, el capítulo 1, donde presentamos una introducción básica del marco general en el que se desarrollan los principales contenidos que aparecen en el texto. Para ello, hacemos una descripción escueta de lo que son los espacios de Banach y de las características de los operadores que se van a tratar. También damos una descripción muy general de los esquemas iterativos en la que se estudia su convergencia, haciendo especial hincapié en el método de Newton, que es el origen de los métodos iterativos que posteriormente se describen y estudian. La aplicación del método de Newton pasa por la evaluación de la derivada primera del operador implicado en cada paso de iteración. Bien porque esta derivada no exista o sea costosa de evaluar, los métodos iterativos que no utilizan derivadas son especialmente interesantes. Un procedimiento para construir métodos iterativos que no utilicen derivadas consiste en aproximar éstas mediante diferencias divididas. De este tipo de métodos nos ocupamos ya en el resto del texto.

La segunda parte del texto centra su atención en una familia de métodos iterativos tipo secante que surge a partir de las interpretaciones geométricas de los métodos de Newton y de la secante, y que son métodos iterativos con memoria. Presentamos un resultado de conver-

gencia semilocal ya conocido, basado en relaciones de recurrencia, vemos cuál es su principal problema y lo resolvemos utilizando dos procedimientos: la construcción, en el capítulo 2, de un método iterativo híbrido (predictor-corrector) y, en el capítulo 3, la modificación de la técnica de demostración de la convergencia presentada en el capítulo 2. Destacamos en el capítulo 3 que la técnica desarrollada para demostrar la convergencia de la familia de métodos tipo secante tiene la ventaja, con respecto a la presentada en el capítulo 2, de que permite tratar ecuaciones en las que el operador implicado puede ser diferenciable o no diferenciable.

La tercera parte del texto está dedicada al método de Steffensen, que es un método iterativo punto a punto que tiene la misma velocidad de convergencia y la misma eficiencia computacional que el método de Newton. Veremos por qué, a pesar de lo anterior, este método es mucho menos utilizado que el método de Newton y qué es lo que podemos hacer para mejorar su utilización. En el capítulo 4, realizamos un análisis desde el punto de vista de la teoría de Kantorovich que nos permite considerar situaciones en las que el operador implicado es diferenciable. En el capítulo 5 lo hacemos desde el punto de vista de una teoría basada en relaciones de recurrencia, que nos permite considerar situaciones en las que el operador implicado, puede se tanto diferenciable como no diferenciable. En ambos capítulos proponemos la utilización de un método iterativo híbrido (predictor-corrector).

Al final del texto hemos añadido algunas referencias bibliográficas que nos han servido de inspiración a lo largo de los últimos años. Por supuesto, ni están todas las que son, ni son todas las que están. Pero seguro que servirán para afianzar el interés de los estudiantes interesados en los contenidos del texto, de manera que este interés les conduzca a una búsqueda bibliográfica más detalla y acorde a sus futuras inquietudes.

No queremos terminar sin dar las gracias a todos aquellos que nos han acompañado a lo largo de todo este tiempo. En particular, a los que, año tras año, se acercan a las *Jornadas de Análisis Numérico y Aplicaciones*, y ya van ocho, que celebramos en la Universidad de La Rioja durante el mes de noviembre. Y, en particular, a nuestros compañeros del grupo de investigación PRIENOL (`https://prienol.unirioja.es`).

Logroño, La Rioja
Julio de 2014

J. A. Ezquerro
M. A. Hernández-Verón

Contenidos

Parte I

PRELIMINARES

En el análisis matemático ordinario trabajamos con el sistema de los números reales o complejos. El análisis funcional se basa en la utilización de espacios lineales, que son generalizaciones de estos sistemas de números. El poder del análisis funcional reside en que permite tratar una gran variedad de problemas que surgen del hecho de que los espacios lineales están compuestos de interesantes objetos matemáticos como son los vectores con un número finito o infinito de componentes o las funciones que satisfacen ciertas condiciones dadas. Muchas transformaciones y ecuaciones importantes del álgebra y del análisis matemático ordinarios se pueden formular en términos de operadores lineales en dichos espacios. El análisis funcional lineal, que es la teoría de los espacios lineales, operadores y ecuaciones, se ha desarrollado como una amplia disciplina matemática con multitud de aplicaciones.

Entre los espacios lineales, los espacios normados completos, que se denominan espacios de Banach, juegan un papel importante. Se denominan así en honor de Stefan Banach y son uno de los objetos de estudio más importantes en el análisis funcional. Los espacios de Banach son típicamente espacios de funciones de dimensión infinita.

Presentamos una introducción básica mínima de esta teoría en esta primera parte del texto con el objetivo de que éste sea autosuficiente.

Por otra parte, es bien conocido que encontrar soluciones de ecuaciones en espacios de Banach, donde el operador implicado es no lineal, es un problema común en las ciencias y la ingeniería. Aunque algunas ecuaciones se pueden resolver analíticamente, generalmente buscamos aproximaciones numéricas de las soluciones, ya que encontrar soluciones exactas habitualmente es difícil. Utilizamos normalmente métodos iterativos para aproximar estas soluciones. Así, introducimos algunas ideas básicas acerca de los métodos iterativos en espacios de Banach y, en particular, del método de Newton, que es el método iterativo más conocido y utilizado para resolver ecuaciones no lineales en espacios de Banach. Finalmente, presentamos algunas ecuaciones no lineales que aparecen frecuentemente en la literatura matemática, que posteriormente utilizamos como aplicaciones para ilustrar los resultados que presentamos en las partes dos y tres del texto.

Capítulo 1

Conceptos generales

El origen de la teoría de los espacios de Banach se encuentra en la publicación del clásico *Théorie des opérations linéaires* [8] por parte del matemático polaco Stefan Banach (1892–1945) en 1932. Desde su nacimiento, esta teoría estuvo relacionada con el resto de ramas del análisis matemático.

Tras unos años de auge, la teoría parecía condenada al ostracismo. Sin embargo, durante los años setenta y principios de los ochenta, la teoría tuvo un nuevo periodo de gran actividad, se resolvieron viejos problemas y, lo que es más importante, se plantearon otros nuevos y se establecieron nuevas conexiones con otras ramas del análisis matemático, tales como análisis armónico, funciones de variable compleja, series ortonormales, teoría de aproximación o probabilidad. De la bibliografía acerca de los espacios de Banach podemos destacar los textos clásicos de Dunford-Schwarz [19] y Day [16], donde se recogen los principales resultados de la época comprendida entre los años treinta y cincuenta. Resultados más recientes pueden encontrarse en los textos de Lindenstrauss-Tzafriri [31],[32] y Beauzamy [10].

Como dice Wojtaszczyk en su introducción [48], no podemos esperar de los espacios de Banach que nos resuelvan los problemas que estemos tratando, pero sí que nos hagan ver dichos problemas con un nuevo enfoque y nos permitan aislar sus características esenciales, obteniendo así una gran generalidad en los resultados. Además, en muchos casos, utilizar las técnicas y los teoremas generales de los espacios de Banach puede sugerirnos nuevos problemas. Al trabajar con espacios de Banach, abarcamos una gran amplitud de situaciones, tales como ecuaciones en el campo real o complejo, sistemas de ecuaciones reales o complejas, ecuaciones diferenciales o ecuaciones integrales.

A continuación, damos una introducción de los conceptos básicos de la teoría de los espacios de Banach que son esenciales en el tratamiento posterior de la resolución, mediante métodos iterativos que utilizan diferencias divididas, de ecuaciones con operadores no lineales en espacios de Banach. Para un estudio más detallado se puede consultar cualquiera de los numerosos tratados que hay en la bibliografía matemática, entre los que citamos los textos de Berberian [11], Rudin [40] y Curtain-Pritchard [15].

Gran cantidad de problemas de las ciencias y la ingeniería se pueden expresar como ecuaciones con operadores no lineales en espacios de Banach. En general, para resolver estas ecuaciones, recurrimos a métodos iterativos. Entre éstos, destaca especialmente, por su simplicidad y eficiencia, el método de Newton. Este método proporciona una herramienta poderosa para la investigación teórica y numérica de las ecuaciones con operadores no lineales. En este capítulo, introducimos la teoría del método de Newton desde dos puntos de

vista: el principio de la mayorante de Kantorovich y una alternativa basada en relaciones de recurrencia, desarrollada por los autores del texto, que proporciona buenos resultados y es sencilla de aplicar.

Terminamos el capítulo presentando algunos tipos importantes de ecuaciones diferenciables e integrales no lineales en espacios de Banach. A continuación, les aplicamos procesos de discretización para transformarlas en sistemas de ecuaciones no lineales que posteriormente resolvemos mediante los métodos iterativos presentados en este texto.

1.1. Espacios de Banach

Los espacios de Banach son una clase importante de espacios lineales normados que reciben su nombre a partir de las contribuciones fundamentales del matemático polaco Stefan Banach al análisis funcional lineal [8].

1.1.1. Espacios lineales normados

En términos abstractos, un espacio vectorial lineal es un conjunto de elementos, llamados habitualmente puntos o vectores, con dos operaciones, la adición y la multiplicación por un escalar, que satisfacen ciertas condiciones. En particular, se requiere que la adición tenga las mismas propiedades algebraicas que la adición de los números reales o complejos.

Cada espacio lineal X tiene asociado un campo escalar $\mathbb{K}$, que es el campo de los números reales $\mathbb{R}$ o el de los complejos $\mathbb{C}$. Llamamos escalar a un número perteneciente a $\mathbb{K}$ y se dice que X es un espacio lineal real o complejo según sea $\mathbb{K} = \mathbb{R}$ o $\mathbb{K} = \mathbb{C}$.

La estructura algebraica de los espacios lineales es similar a la del sistema de los números reales o complejos, permitiendo extender muchas técnicas algebraicas a problemas en un marco más general. Sin embargo, al tratar con otros conceptos de importancia teórica y computacional, como la exactitud de aproximación o la convergencia de sucesiones o series, es necesario introducir una estructura adicional en estos espacios. La estructura métrica (o topológica) que consideramos aquí se basa en una simple generalización de la idea de valor absoluto de un número real o módulo de un número complejo.

Supongamos que, para cada elemento x de un espacio lineal X, se define norma de x como el número real no negativo $\|x\|$ que satisface las siguientes tres condiciones:

N_1: $\|x\| > 0$ si $x \neq 0$, $\|0\| = 0$,

N_2: $\|\lambda x\| = |\lambda|\|x\|$ para todo $\lambda \in \mathbb{K}$,

N_3: $\|x + y\| \leq \|x\| + \|y\|$, para todo $y \in X$.

Entonces, un espacio lineal X en el que hay definida una norma se llama espacio lineal normado.

Es fácil ver que $\mathbb{R}$ y $\mathbb{C}$ son espacios lineales normados con $\|x\| = |x|$. Desde un punto de vista geométrico, podemos interpretar $\|x\|$ como la distancia del origen 0 del espacio al punto x o como la longitud del vector x.

La distancia $d(x, y)$ de un punto x a un punto y se define en un espacio lineal normado X como $d(x, y) = \|x - y\|$. Por tanto, si consideramos y como una aproximación a x, el error de la aproximación es $\|x - y\|$.

En general, en un espacio lineal se pueden introducir diferentes normas. Por ejemplo, en $\mathbb{R}^n$ y $\mathbb{C}^n$ se definen

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}, \quad 1 \leq p < \infty, \tag{1.1}$$

y

$$\|x\|_\infty = \max_{(i)} |x_i|, \quad i = 1, 2, \ldots, n, \tag{1.2}$$

donde $x = (x_1, x_2, \ldots, x_n)$. Obviamente estas definiciones satisfacen las condiciones N_1 y N_2. Para verificar N_3 se utiliza la desigualdad de Minkowski [29]:

$$\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}}, \qquad 1 \leq p \leq \infty.$$

La norma infinito es el caso límite cuando $p \to \infty$. Para verlo, suponemos que $\|x\|_\infty = |x_1| \neq 0$ y escribimos

$$\|x\|_p = |x_1| \left(1 + \sum_{i=2}^{n} \left|\frac{x_i}{x_1}\right|^p\right)^{\frac{1}{p}}.$$

Como $\left|\dfrac{x_i}{x_1}\right| \leq 1$, entonces

$$\lim_{p\to\infty} \|x\|_p = |x_1| = \|x\|_\infty.$$

Los espacios $\mathbb{R}^n$ y $\mathbb{C}^n$ con la norma $\|x\|_p$ se denotan, respectivamente, por $\mathbb{R}^n_p$ y $\mathbb{C}^n_p$. Designamos con $\mathbb{R}^\infty$ (o $\mathbb{C}^\infty$) al conjunto de sucesiones de números reales (o complejos). En $\mathbb{R}^\infty$, por ejemplo, para $p = 1$ la serie infinita

$$\|x\|_1 = \sum_{i=1}^{\infty} |x_i|$$

convergerá en un subconjunto de $\mathbb{R}^\infty$. Este subconjunto es un subespacio normado de $\mathbb{R}^\infty$ que se denota por ℓ^1.

De la misma forma, para $1 \leq p < \infty$, podemos definir ℓ^p como el subconjunto de $\mathbb{R}^\infty$ (o $\mathbb{C}^\infty$) formado por los vectores x tales que

$$\|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{\frac{1}{p}} < \infty.$$

Para ver que estos subconjuntos son espacios normados, debemos considerar la desigualdad de Minkowski para series infinitas [29],

$$\left(\sum_{i=1}^{\infty} |x_i + y_i|^p\right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{\infty} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{\infty} |y_i|^p\right)^{\frac{1}{p}}, \quad 1 \leq p < \infty,$$

y comprobar las condiciones de subespacio.

El espacio ℓ^∞ es el conjunto de todas las sucesiones acotadas reales (o complejas) con la norma definida en (1.2):

$$\|x\|_\infty = \sup_{(i)}\{|x_i|\}, \quad x = (x_1, x_2, \ldots).$$

Se verifica la siguiente relación de contenido [40]:

$$\ell^1 \subseteq \ell^2 \subseteq \cdots \subseteq \ell^\infty.$$

Es natural, para espacios de funciones, reemplazar el sumatorio en la definición de norma por una integral. Por ejemplo, el subconjunto $L^1([0,1])$ del espacio $F([0,1])$ de todas las funciones reales $x = x(t)$, con $0 \leq t \leq 1$, para las cuales la integral de Riemann

$$\|x\|_1 = \int_0^1 |x(t)|\, dt$$

existe y es finita, es un espacio normado, como lo son los espacios $L^p([0,1])$ de las funciones reales para los cuales las integrales

$$\|x\|_p = \left(\int_0^1 |x(t)|^p\, dt\right)^{\frac{1}{p}}, \qquad 1 \leq p < \infty,$$

existen y son finitas. El espacio $L^\infty([0,1])$ está formado por todas las funciones reales acotadas con la norma

$$\|x\|_\infty = \sup_{t\in[0,1]} |x(t)|.$$

Los conjuntos $L^p([0,1])$ cumplen la siguiente relación [40]

$$L^1([0,1]) \supseteq L^2([0,1]) \supseteq \cdots \supseteq L^\infty([0,1]).$$

Indicamos, a continuación, cómo relacionar las diferentes normas que podemos definir en un espacio lineal. El concepto que las relaciona es el de normas equivalentes. Así, se dice que dos normas diferentes $\|x\|$ y $\|x\|'$ de un espacio lineal son equivalentes si existen constantes a, b tales que $0 < a < b$ y

$$a\|x\| \leq \|x\|' \leq b\|x\|, \quad \text{para todo} \quad x \in X.$$

Es conocido que todas las normas son equivalentes en un espacio lineal de dimensión finita [29].

1.1.2. Operadores lineales y acotados en espacios de Banach. Inversión de operadores

En un espacio lineal normado X podemos definir las nociones analíticas de convergencia y límite de una sucesión de elementos del espacio. Denominamos límite de una sucesión $\{x_n\}$ de elementos de X a un elemento x^* tal que

$$\lim_{n\to\infty} \|x_n - x^*\| = 0 \tag{1.3}$$

es decir: si, dado $\varepsilon > 0$, existe $N \in \mathbb{N}$ tal que $\|x^* - x_n\| < \varepsilon$ para $n > N$. Si se cumple (1.3), decimos entonces que la sucesión $\{x_n\}$ converge a x^* y denotamos $\lim_{n\to\infty} x_n = x^*$.

Una serie infinita $x_1 + x_2 + x_3 + \cdots$ de elementos de un espacio lineal normado es convergente si la sucesión $\{x_n\}$ de sumas parciales

$$x_n = \sum_{i=1}^n x_i$$

converge. En este caso, denominamos suma de la serie infinta al límite x^* de $\{x_n\}$.

Una clase importante de los espacios lineales normados son los llamados espacios de Banach. El asunto a tratar en la resolución de muchos problemas es la existencia de un límite x^* de una sucesión $\{x_n\}$ en un espacio lineal normado X; situación común en el análisis matemático clásico vinculada a la idea de completitud.

Para introducir la propiedad anterior veamos un ejemplo ilustrativo. Consideramos la sucesión $\{q_n\}$ de números racionales definida por

$$q_0 = 1, \quad q_{n+1} = \frac{1}{2}\left(q_n + \frac{3}{q_n}\right), \quad n = 0, 1, \ldots \tag{1.4}$$

Es conocido que no existe ningún número racional q^* que pueda interpretarse como límite de esta sucesión. Sin embargo, si estudiamos la sucesión anterior en el ámbito de los números reales, vemos que tiene límite q^* y que además es solución de la ecuación $x^2 = 3$, [38]. Al generalizar esta propiedad que poseen los números reales a otros espacios aparece la idea de completitud. Es interesante destacar que la completitud no se deduce de las propiedades de los espacios normados, sino que es una propiedad adicional que un espacio puede tener o no.

Definición 1.1. *Una sucesión $\{x_n\}$ de elementos de un espacio lineal normado se dice de Cauchy si*

$$\lim_{n\to\infty} \lim_{m\to\infty} \|x_{n+m} - x_n\| = 0.$$

La elección de una norma puede ser de suma importancia cuando tratamos con procesos infinitos en espacios de dimensión infinita.

Definición 1.2. *Se dice que un espacio normado X es completo si toda sucesión de Cauchy es convergente a un límite que es un elemento de X.*

Así, la sucesión de números racionales definida por (1.4) es una sucesión de Cauchy, pero no tiene un límite racional. En consecuencia, el conjunto $\mathbb{Q}$ de los números racionales no es completo. Sin embargo, la sucesión anterior vista como una sucesión de números reales sí que tiene límite. De hecho, el conjunto $\mathbb{R}$ de los números reales es completo.

Definición 1.3. *Un espacio de Banach es un espacio normado completo.*

A continuación, damos algunos ejemplos de espacios de Banach.

Ejemplo 1.4. Los espacios $\mathbb{R}^n$ y $\mathbb{C}^n$ con las normas $\|\cdot\|_p$, $1 \leq p \leq \infty$, definidas en (1.1) y (1.2) son espacios de Banach.

Demostración. Si un espacio normado de dimensión finita es completo para una norma, es completo para cualquier otra norma equivalente. Por consiguiente, será suficiente probarlo para la norma $\|\cdot\|_\infty$.

Si $\{x_m = (x_{m1}, x_{m2}, \ldots, x_{mn})\}$ es una sucesión de Cauchy en uno de estos espacios, entonces las sucesiones $\{x_{mi}\}$, $i = 1, \ldots, n$, son de Cauchy en $\mathbb{R}$ o $\mathbb{C}$.

Por ser estos espacios completos, existen los números x_i^*, $i = 1, 2, \ldots, n$, tales que

$$\lim_{m\to\infty} |x_{mi} - x_i^*| = 0, \quad i = 1, 2, \ldots, n.$$

Se cumple entonces

$$\{x_{mi}\} \to x_i^*, \quad 1 \leq i \leq n, \quad \text{si y sólo si} \quad \{x_m\} \to x^*,$$

con $x^* = (x_1^*, x_2^*, \dots, x_n^*)$. Por tanto, la sucesión $\{x_m\}$ tiene por límite x^*. Luego, tanto $\mathbb{R}^n$ como $\mathbb{C}^n$ son completos. ■

También podemos probar, de la misma forma que en el ejemplo 1.4, que ℓ^p, con $1 \leq p \leq \infty$, es un espacio de Banach.

Si ahora pensamos en los espacios de funciones y, por tanto, en sucesiones de funciones, la idea de convergencia de una sucesión de funciones admite ciertos matices.

Definición 1.5. *Sean un subconjunto S del espacio normado X, una sucesión de funciones $\{f_n\}_{n\geq 1}$ tales que $f_n : S \longrightarrow X$ y una función $f : S \longrightarrow X$. Decimos que la sucesión $\{f_n\}_{n\geq 1}$ converge puntualmente a f en S si, para $s \in S$, la sucesión $\{f_n(s)\}_{n\geq 1}$ converge a $f(s)$; es decir: fijado $s \in S$, para todo $\varepsilon > 0$, existe $N \in \mathbb{N}$ tal que $\|f(s) - f_n(s)\| < \varepsilon$ para $n > N$.*

Es fundamental observar en la definición anterior que la elección de N se hace después de conocer s y ε, de modo que N depende de ambos valores.

Notemos que si bien en los espacios de dimensión finita la elección de una norma no modifica el carácter de completitud, ya que todas las normas son equivalentes ([40]), en espacios de dimensión no finita sí puede determinar la completitud del espacio. Veamos a continuación, como ejemplo, que el espacio $\mathcal{C}([a,b])$ de las funciones reales de una variable que son continuas en el intervalo cerrado $[a,b]$ no es un espacio completo con la norma

$$\|x\| = \left(\int_a^b |x(t)|^2 \, dt \right)^{\frac{1}{2}}.$$

Por ejemplo, si $[a,b] = [0,1]$, la sucesión de funciones continuas $\{x_m(t)\}$ definida por

$$x_m(t) = \begin{cases} 2^m t^{m+1}, & 0 \leq t \leq \dfrac{1}{2}, \\ 1 - 2^m (1-t)^{m+1}, & \dfrac{1}{2} \leq t \leq 1, \end{cases}$$

es una sucesión de Cauchy con la norma anterior (figura 1.1) y tiene por límite la sucesión (figura 1.2)

$$x^*(t) = \begin{cases} 0, & 0 \leq t < \dfrac{1}{2}, \\ \dfrac{1}{2}, & t = \dfrac{1}{2}, \\ 1, & \dfrac{1}{2} < t \leq 1, \end{cases}$$

pero $x^*(t) \notin \mathcal{C}([0,1])$.

Acabamos de ver que el límite de una sucesión de funciones continuas tiene como límite una función que no lo es, lo que implica que el espacio $\mathcal{C}([a,b])$ no sea completo con la norma utilizada. La idea de que la función límite conserve las propiedades que tiene el conjunto de funciones que definen la sucesión pasa por un concepto de convergencia más sutil, el de convergencia uniforme que definimos a continuación.

Definición 1.6. *Sean un subconjunto S del espacio normado X, una sucesión de funciones $\{f_n\}_{n\geq 1}$ tales que $f_n : S \longrightarrow X$ y una función $f : S \longrightarrow X$. Decimos que la sucesión $\{f_n\}_{n\geq 1}$ converge uniformemente a f en S si, para cada $\varepsilon > 0$, existe $N \in \mathbb{N}$ tal que $\|f(s) - f_n(s)\| < \varepsilon$ para $n > N$ y $s \in S$.*

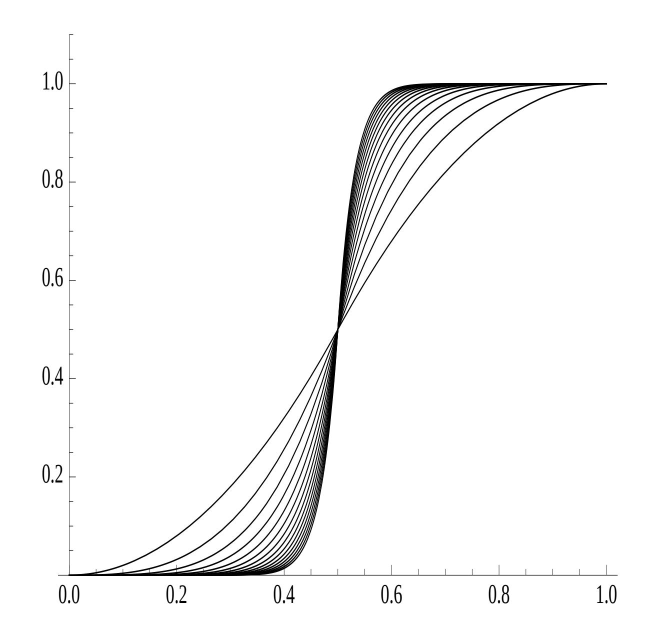


Figura 1.1: Representación gráfica de $x_m(t)$.

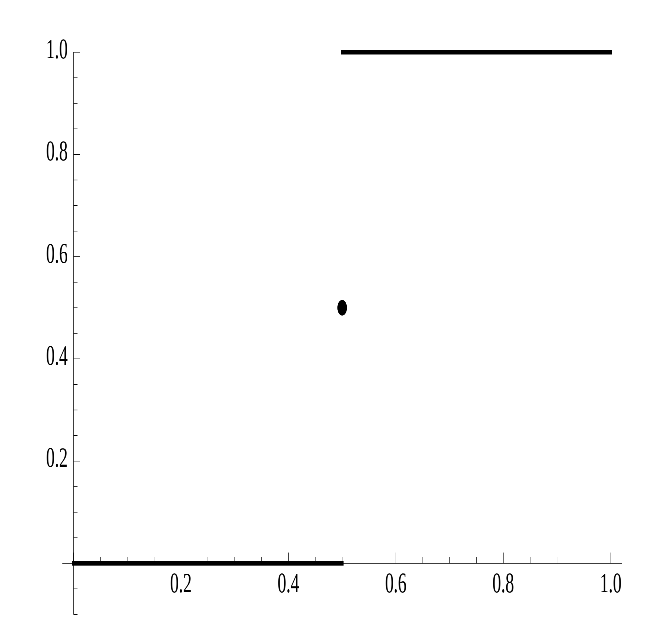


Figura 1.2: Representación gráfica de $x^*(t)$.

Destacamos que en la definición anterior el valor de N es el mismo para todos los elementos $s \in S$, mientras que, en la convergencia puntual, N depende de ε y $s \in S$. Este hecho va a permitir que la función límite uniforme herede propiedades de las funciones que forman la sucesión. Así, por ejemplo, si $\{f_n\}_{n\geq 1}$ es una sucesión de funciones continuas con límite uniforme f, entonces f es una función continua. En el siguiente ejemplo utilizamos lo anterior para ver que $\mathcal{C}([a,b])$ es un espacio de Banach con la norma del máximo.

Ejemplo 1.7. El espacio $\mathcal{C}([a,b])$ es un espacio de Banach con la norma

$$\|x\|_\infty = \max_{t\in[a,b]} |x(t)|.$$

DEMOSTRACIÓN. Sea $\{x_m\}$ una sucesión de Cauchy en $\mathcal{C}([a,b])$. Entonces, dado $\varepsilon > 0$, existe un número natural N tal que, para todo m y n tales que $m, n > N$, se tiene

$$\|x_m - x_n\|_\infty = \max_{t\in[a,b]} |x_m(t) - x_n(t)| < \varepsilon. \tag{1.5}$$

Fijado cualquier $t_0 \in [a,b]$,

$$|x_m(t_0) - x_n(t_0)| < \varepsilon, \quad m, n > N,$$

y $\{x_1(t_0), x_2(t_0), \ldots\}$ es una sucesión de Cauchy de números reales. Como $\mathbb{R}$ es completo, $x_m(t_0) \to x(t_0)$ cuando $m \to \infty$. De esta forma, podemos asociar a cada $t \in [a,b]$ un único $x(t)$. Esto define una función x en $[a,b]$. Veamos que $x \in \mathcal{C}([a,b])$ y $x_m \to x$.

De (1.5), cuando $n \to \infty$, se tiene

$$\max_{t\in[a,b]} |x_m(t) - x(t)| \leq \varepsilon, \quad m > N.$$

Por tanto, para cada $t \in [a,b]$, tenemos

$$|x_m(t) - x(t)| \leq \varepsilon, \quad m > N,$$

lo que implica que $\{x_m\}$ converge a x uniformemente en $[a,b]$. Como las funciones x_m son continuas y la convergencia es uniforme, la función límite x es continua en $[a,b]$. $\blacksquare$

Ejemplo 1.8. También se puede ver que los espacios $\mathcal{C}^n([a,b])$ son espacios de Banach con las normas
$$\|x\|_\infty = \max\{\|x\|, \|x'\|, \ldots, \|x^{(n)}\|\}$$
y
$$\|x\|_1 = \|x\| + \|x'\| + \cdots + \|x^{(n)}\|,$$
donde $\|\cdot\|$ denota la norma en $\mathcal{C}([a,b])$, [8].

En cálculo se trabaja con funciones reales definidas sobre un intervalo real. En análisis funcional, consideramos espacios más generales, tales como espacios métricos o normados y aplicaciones entre dichos espacios. En este caso, a estas aplicaciones se les llama operadores. Diremos que un operador T aplica el espacio X en Y si a un elemento $x \in X$ le hace corresponder un elemento $T(x) \in Y$. En principio, T no tiene por qué estar definido sobre todo el espacio X ni recorrer todos los valores de Y. Al conjunto de puntos de X donde está definido T lo llamamos dominio de T, y lo denotaremos por $\mathcal{D}(T)$, y al conjunto de valores de Y que toma el operador T lo llamamos rango de T, y lo denotamos por $\mathcal{R}(T)$.

De especial interés son aquellos operadores que conservan las operaciones algebraicas de los espacios donde están definidos.

Definición 1.9. *Un operador L que aplica un espacio lineal X en otro espacio lineal Y, ambos espacios sobre el mismo cuerpo $\mathbb{K}$, se dice lineal si, para todo $x_1, x_2 \in X$, tenemos*
$$L(x_1 + x_2) = L(x_1) + L(x_2),$$
y, para todo $x \in X$ y todo $\lambda \in \mathbb{K}$,
$$L(\lambda x) = \lambda L(x).$$

Es muy corriente en análisis funcional emplear la notación Lx en lugar de $L(x)$ para los operadores lineales.

Definición 1.10. *Un operador T entre dos espacios normados X e Y se dice acotado si existe un número real c tal que, para todo $x \in \mathcal{D}(T)$, se cumple*
$$\|T(x)\| \leq c\|x\|. \tag{1.6}$$

Observamos que la norma de la izquierda es la del espacio Y y la de la derecha la del espacio X, aunque las hayamos denotado igual.

Notemos que en el caso de operadores acotados, la imagen no tiene por qué estar acotada, lo que verifican estos operadores es que transforman conjuntos acotados en conjuntos acotados.

Si X e Y son dos espacios lineales sobre un mismo cuerpo $\mathbb{K}$, entonces el conjunto de todos los operadores lineales y acotados entre X e Y, que denotamos a partir de ahora por $\mathcal{L}(X,Y)$, es un espacio lineal sobre el cuerpo $\mathbb{K}$ con las operaciones
$$(L_1 + L_2)(x) = L_1x + L_2x, \quad \text{para todo } x \in X,$$
$$(\lambda L)x = \lambda(Lx), \quad \text{para todo } x \in X \text{ y todo } \lambda \in \mathbb{K}.$$
A continuación, vamos a definir una norma en el conjunto de operadores acotados. Dejando a un lado el caso en que $x = 0$, nos podemos preguntar cuál es el menor número c que verifica (1.6). De aquí surge la idea para definir la norma de un operador.

Definición 1.11. *Sea T un operador acotado. Definimos la norma de T como*

$$\|T\| = \text{ínf}\{c; \|T(x)\| \leq c\|x\|,\ x \in \mathcal{D}(T),\ x \neq 0\} = \sup_{\substack{x \in \mathcal{D}(T) \\ x \neq 0}} \frac{\|T(x)\|}{\|x\|}.$$

Como consecuencia, si T es acotado, entonces

$$\|T(x)\| \leq \|T\|\|x\|. \tag{1.7}$$

Otra fórmula alternativa para la norma anterior, cuando el operador T es lineal, es la siguiente:

$$\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ \|x\|=1}} \|T(x)\|. \tag{1.8}$$

Que estas fórmulas son equivalentes y que todas ellas son normas puede verse en [29]. Además, el espacio $\mathcal{L}(X, Y)$ con Y completo y la norma (1.8) es un espacio completo [29].

Antes de considerar propiedades generales de los operadores lineales acotados veamos algunos ejemplos típicos.

Ejemplo 1.12. El operador identidad, $I : X \to X$, definido por $Ix = x$, es un operador acotado y su norma es $\|I\| = 1$.

Ejemplo 1.13. El operador cero, $0 : X \to X$, definido por $0x = 0$ para todo $x \in X$, es un operador acotado y su norma es $\|0\| = 0$.

Ejemplo 1.14. Sea X el espacio normado de todos los polinomios en $J = [0, 1]$ con la norma $\|x\| = \text{máx}\,|x(t)|$, $t \in J$. El operador diferenciación T, definido en X por

$$Tx(t) = x'(t),$$

es lineal, pero no acotado. En efecto, sea $x_n(t) = t^n$ con $n \in \mathbb{N}$. Entonces $\|x_n\| = 1$ y

$$Tx_n(t) = x_n'(t) = nt^{n-1}.$$

Por tanto, $\|Tx_n\| = n$ y $\dfrac{\|Tx_n\|}{\|x_n\|} = n$. Como $n \in \mathbb{N}$, no existe un número fijo c tal que $\dfrac{\|Tx_n\|}{\|x_n\|} \leq c$. A partir de lo anterior y de (1.6), se concluye que T no es acotado.

Teorema 1.15. *([29]) Si un espacio normado X es de dimensión finita, entonces cada operador lineal en X es acotado.*

Ejemplo 1.16. Sean $X = \mathbb{R}^m$ e $Y = \mathbb{R}^n$. Como X e Y son espacios de dimensión finita, el conjunto de operadores lineales de X en Y es $\mathcal{L}(X, Y)$ y, por tanto, existe una correspondencia biyectiva entre $\mathcal{L}(X, Y)$ y el conjunto $M_{n\times m}(\mathbb{R})$ de las matrices reales de n filas y m columnas de la siguiente manera.

Si $L \in \mathcal{L}(X, Y)$ y $A = (a_{ij})$, $1 \leq i \leq n$, $1 \leq j \leq m$, es su matriz asociada, entonces

$$y = Lx = Ax, \quad \text{para todo} \quad x \in \mathcal{D}(X),$$

donde $x = (x_1, x_2, \dots, x_m) \in X$, $y = (y_1, y_2, \dots, y_n) \in Y$ e

$$y_j = \sum_{k=1}^{m} \alpha_{jk} x_k, \; j = 1, 2, \dots, n.$$

Además, la norma de $\|L\|$ viene definida por

$$\|L\| = \sup_{\substack{x \in \mathcal{D}(L) \\ x \neq 0}} \frac{\|Lx\|_Y}{\|x\|_X} = \sup_{\substack{x \in \mathcal{D}(L) \\ x \neq 0}} \frac{\|Ax\|_Y}{\|x\|_X}.$$

En el espacio de las matrices de n filas y m columnas podemos definir varias normas. Si denotamos la norma en el espacio X por $\|\cdot\|_X$ y la norma en el espacio Y por $\|\cdot\|_Y$, diremos que una norma $\|\cdot\|$ en el espacio de las matrices es compatible con $\|\cdot\|_X$ y $\|\cdot\|_Y$ si

$$\|Ax\|_Y \leq \|A\| \|x\|_X.$$

Se cumple entonces que la norma definida por

$$\|A\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|Ax\|_Y}{\|x\|_X}$$

es compatible con $\|\cdot\|_X$ y $\|\cdot\|_Y$. Esta norma se denomina a menudo norma natural definida por $\|\cdot\|_X$ y $\|\cdot\|_Y$.

Si consideramos en X e Y la norma infinito definida en (1.2), se tiene que la norma natural es

$$\|A\| = \max_i \sum_{j=1}^{m} |a_{ij}|.$$

Definición 1.17. *Un operador T entre dos espacios normados X e Y se dice continuo en un punto $\tilde{x} \in X$ si, para toda sucesión $\{x_n\}$ tal que $\lim_{n\to\infty} x_n = \tilde{x}$, se tiene $\lim_{n\to\infty} T(x_n) = T(\tilde{x})$. Además, diremos que un operador es continuo si lo es en todos los puntos de su dominio.*

La definición anterior generaliza la idea de función continua del cálculo en una variable. Los operadores lineales tienen la siguiente propiedad.

Teorema 1.18. *Sea T un operador lineal entre dos espacios normados. Entonces,*

(*i*) *T es continuo si y sólo si es acotado.*

(*ii*) *Si T es continuo en un punto, entonces T es continuo en todos los puntos de su dominio.*

Demostración. Comenzamos probando el apartado (i). En primer lugar, probamos que si T es acotado, entonces es continuo. Para $T = 0$ resulta trivial. Si $T \neq 0$, entonces $\|T\| \neq 0$. Suponemos que T es acotado y $\tilde{x} \in \mathcal{D}(T)$. Dado $\varepsilon > 0$, para cada $x \in \mathcal{D}(T)$ tal que

$$\|x - \tilde{x}\| < \delta, \quad \text{donde} \quad \delta = \frac{\varepsilon}{\|T\|},$$

obtenemos, al ser T lineal,

$$\|Tx - T\tilde{x}\| = \|T(x - \tilde{x})\| \leq \|T\| \, \|x - \tilde{x}\| < \|T\|\delta = \varepsilon.$$

Como $\tilde{x} \in \mathcal{D}(T)$, entonces T es continuo.

Recíprocamente, suponemos que T es continuo y $\tilde{x} \in \mathcal{D}(T)$. Entonces, dado $\varepsilon > 0$, existe $\delta > 0$ tal que

$$\|Tx - T\tilde{x}\| < \varepsilon, \text{ para todo } x \in \mathcal{D}(T), \text{ cumpliendo } \|x - \tilde{x}\| < \delta.$$

Sea $y \in \mathcal{D}(T)$ tal que $y \neq 0$. Consideramos $|\theta| < \delta$ y $x = \tilde{x} + \dfrac{\theta}{\|y\|}y$ perteneciente a $\mathcal{D}(T)$ por ser T lineal. Como $x - \tilde{x} = \dfrac{\theta}{\|y\|}$, entonces $\|x - \tilde{x}\| < \delta$. Aplicando (1.7), por ser T lineal, obtenemos

$$\|Tx - T\tilde{x}\| = \|T(x - \tilde{x})\| = \left\|T\left(\frac{\theta}{\|y\|}y\right)\right\| = \frac{|\theta|}{\|y\|}\|Ty\|,$$

lo que implica

$$\frac{|\theta|}{\|y\|}\|Ty\| < \varepsilon \qquad \text{y} \qquad \|Ty\| < \frac{\varepsilon}{|\theta|}\|y\|.$$

Podemos escribir entonces $\|Ty\| \leq c\|y\|$, donde $c = \dfrac{\varepsilon}{|\theta|}$. Y, como para $y = 0$ resulta trivial, T es acotado.

Veamos ahora el apartado (ii). Si T es continuo en un punto, por la segunda parte de la demostración de (i), T es acotado, lo que implica, de nuevo por (i), que T es continuo en todos sus puntos. ■

Como consecuencia del teorema 1.18, a los operadores lineales continuos entre espacios normados se les suele llamar también operadores lineales acotados.

A continuación, introducimos los conceptos de Lipschitz continuidad y Hölder continuidad para un operador.

Definición 1.19. *Un operador T entre dos espacios normados X e Y se dice que satisface una condición Lipschitz (o que es Lipschitz continuo) si existe una constante $C \geq 0$ tal que*

$$\|T(x) - T(y)\| \leq C\,\|x - y\|, \quad \forall x, y \in X.$$

En tal caso, C se llama la constante de Lipschitz del operador T.

Notemos que si un operador satisface una condición de Lipschitz, en particular es continuo.

Una generalización conocida de la propiedad de que un operador sea Lipschitz continuo es que sea Hölder continuo.

Definición 1.20. *Un operador T entre dos espacios normados X e Y se dice que satisface una condición (C,p)-Hölder (o que es (C,p)-Hölder continuo) si existen dos constantes $C \geq 0$ y $p \in [0,1]$ tales que*

$$\|T(x) - T(y)\| \leq C\,\|x - y\|^p, \quad \forall x, y \in X.$$

En tal caso, p se llama exponente de la condición de Hölder. Observamos que si $p = 1$, entonces el operador anterior T satisface una condición Lipschitz.

Ahora introducimos el operador inverso de uno dado, operador fundamental para resolver ecuaciones de la forma $Lx = y$, donde L es un operador lineal.

Definición 1.21. *Sea L es un operador lineal y acotado de X en Y. Si existe un operador L^{-1} que aplica $\mathcal{R}(L)$ en $\mathcal{D}(L)$, de manera que*

$$L^{-1}Lx = x, \quad \text{para todo } x \in \mathcal{D}(L),$$

$$LL^{-1}y = y, \quad \text{para todo } y \in \mathcal{R}(L),$$

diremos que L tiene inverso y que L^{-1} es el inverso de L.

Se puede probar que si L^{-1} existe, entonces también es un operador lineal. Una propiedad algebraica que garantiza la existencia del inverso es la siguiente: si de la relación $Lx = 0$ se deduce que x debe ser cero, entonces L tiene inverso. Esta condición es además una condición necesaria [38].

La condición anterior es en general difícil de comprobar. Necesitaremos entonces una caracterización analítica del concepto de inversión. Para ello, los siguientes resultados son fundamentales [27],[38].

Lema 1.22. (Lema de Banach) *Sea L un operador lineal acotado en un espacio de Banach X verificando*

$$\|L\| \leq k < 1.$$

Entonces, el operador $I - L$ tiene inverso continuo y cumple

$$\|(I - L)^{-1}\| \leq \frac{1}{1-k}.$$

Demostración. Si L_1 y L_2 son operadores en X, entonces, por (1.7), se cumple que $\|L_1L_2\| \leq \|L_1\|\|L_2\|$. Aplicando reiteradamente lo anterior, obtenemos

$$\|L^n\| \leq \|L\|^n, \quad n = 1, 2, \dots \tag{1.9}$$

Consideramos ahora la serie

$$S = I + L + L^2 + \cdots + L^n + \cdots ,$$

que converge, ya que $\|L\| < 1$ y está mayorada por la serie numérica convergente

$$1 + \|L\| + \|L\|^2 + \cdots + \|L\|^n + \cdots$$

Como $\mathcal{L}(X, X)$ es completo, S converge y es un operador de $\mathcal{L}(X, X)$. Además,

$$S(I - L) = (I + L + \cdots + L^n + \cdots)(I - L) = I$$

y, análogamente, $(I - L)S = I$. Por tanto, $S = (I - L)^{-1}$. Finalmente, por (1.9),

$$\|S\| \leq \|I\| + \|L\| + \cdots + \|L^n\| + \cdots \leq 1 + k + \cdots + k^n + \cdots = \frac{1}{1-k}$$

que da la cota para $(I - L)^{-1}$. ■

Unas ligeras variantes del lema anterior son los siguientes resultados.

Lema 1.23. *Sea L un operador lineal acotado en un espacio de Banach X. Existe L^{-1} si y sólo si existe un operador inversible M en X tal que*

$$\|I - ML\| < 1.$$

En este caso,

$$L^{-1} = \sum_{n=0}^{\infty} (I - ML)^n M \qquad y \qquad \|L^{-1}\| \leq \frac{\|M\|}{1 - \|I - ML\|}.$$

Demostración. Tomando $I - ML$ en lugar de L en el lema 1.22 se asegura la existencia del operador continuo inverso de $I - (I - ML) = ML$, que está dado por

$$(ML)^{-1} = \sum_{n=0}^{\infty} (I - ML)^n.$$

Por ser M inversible, se tiene

$$L^{-1} = \sum_{n=0}^{\infty} (I - ML)^n M.$$

Por otro lado, teniendo que

$$\|(ML)^{-1}\| \leq \frac{1}{1 - \|I - ML\|},$$

obtenemos

$$\|L^{-1}\| = \|L^{-1}M^{-1}M\| = \|(ML)^{-1}M\| \leq \|(ML)^{-1}\| \, \|M\| = \frac{\|M\|}{1 - \|I - ML\|}.$$

Finalmente la condición necesaria se prueba tomando $M = L^{-1}$, ya que

$$\|I - ML\| = \|I - I\| = 0 < 1. \quad \blacksquare$$

Lema 1.24. *Sea L un operador lineal acotado en un espacio de Banach X. Existe L^{-1} si y sólo si existe un operador lineal acotado M en X tal que existe M^{-1} y*

$$\|M - L\| < \frac{1}{\|M^{-1}\|}.$$

En este caso,

$$L^{-1} = \sum_{n=0}^{\infty} (I - M^{-1}L)^n M^{-1}$$

y

$$\|L^{-1}\| \leq \frac{\|M^{-1}\|}{1 - \|I - M^{-1}L\|} \leq \frac{\|M^{-1}\|}{1 - \|M^{-1}\|\|M - L\|}.$$

Demostración. Para la condición suficiente tomamos M^{-1} en lugar de M en el lema 1.23, cumpliéndose entonces

$$\|I - M^{-1}L\| = \|M^{-1}(M - L)\| \leq \|M^{-1}\|\|M - L\| < 1.$$

La condición necesaria se prueba tomando $M = L$. $\blacksquare$

1.2. Diferenciación de operadores

En esta sección presentamos algunos de los conceptos y resultados básicos del cálculo diferencial en espacios de Banach. A parte del interés que esta sección pueda tener por sí misma, se establece en ella el contexto que permitirá estudiar la aproximación de soluciones de ecuaciones con operadores no lineales mediante métodos iterativos.

Comenzamos dando una primera definición de derivada para operadores definidos en espacios de Banach.

Definición 1.25. *Dado $x_0 \in X$, si existe un operador $L \in \mathcal{L}(X,Y)$ que, para todo $x \in X$, cumple*

$$\lim_{h \to 0} \frac{F(x_0 + hx) - F(x_0)}{h} = L(x), \tag{1.10}$$

se dice entonces que F es diferenciable Gateaux (o diferenciable débilmente) en x_0. En esta situación, el operador lineal L es la derivada Gateaux de F en x_0 y denotamos $L = F'(x_0)$.

Sin embargo, la derivada Gateaux no es una buena generalización del concepto de derivada para funciones escalares, como prueba el hecho de que existan funciones derivables Gateaux y no continuas. Por ejemplo, la función $F : \mathbb{R}^2 \to \mathbb{R}$ definida por

$$F(x,y) = \begin{cases} \dfrac{xy^2}{x^2 + y^4}, & \text{si } (x,y) \neq (0,0), \\ 0, & \text{si } (x,y) = (0,0), \end{cases}$$

es derivable Gateaux en el punto $(0,0)$ y, sin embargo, no es continua en dicho punto. Para estar seguros de que las funciones diferenciables son continuas, introducimos el siguiente concepto más fuerte de derivada.

Definición 1.26. *Si el límite de la ecuación (1.10) es uniforme en el conjunto $\{x \in X; \|x\| = 1\}$, entonces se dice que F es diferenciable Fréchet (o simplemente diferenciable) en x_0. En este caso, el operador lineal $L = F'(x_0)$ se llama derivada de F en x_0.*

Equivalentemente, el concepto de diferenciabilidad se puede expresar de la siguiente forma. Dado $x_0 \in X$, si existe un operador lineal y continuo L de X en Y de manera que

$$\lim_{\|v\| \to 0} \frac{\|F(x_0 + v) - F(x_0) - Lv\|}{\|v\|} = 0,$$

entonces F es diferenciable Fréchet en x_0 y el operador $F'(x_0) = L$ se denomina derivada de F en x_0.

Como consecuencia de la definición anterior, tenemos las siguientes propiedades de la derivada, cuyas demostraciones pueden consultarse en [27]:

(*i*) Si F es diferenciable Fréchet en x_0, entonces F es diferenciable Gateaux en x_0.

(*ii*) Si un operador es diferenciable en un punto x_0, entonces es continuo en dicho punto.

(*iii*) Si L es un operador lineal de X en Y, entonces $L'(x) = L$ para todo $x \in X$.

(*iv*) Si P y Q son operadores de X en Y, la suma $P+Q$ es el operador definido por $(P+Q)(x) = P(x) + Q(x)$, $x \in X$. Si P y Q son diferenciables en x_0, entonces

$$(P+Q)'(x_0) = P'(x_0) + Q'(x_0).$$

(*v*) Sean X, Y y Z espacios de Banach, Q un operador de X en Z y P un operador de Z en Y. La composición PQ es un operador de X en Y definido por

$$PQ(x) = P\left(Q(x)\right), \ x \in X.$$

Si Q es diferenciable en x_0 y P es diferenciable en $z_0 = Q(x_0)$, entonces PQ es diferenciable en x_0 y

$$(PQ)'(x_0) = P'\left(Q(x_0)\right)Q'(x_0) = P'(z_0)Q'(x_0).$$

(*vi*) Combinando (*iii*) y (*v*) tenemos que si L es un operador lineal acotado de Z en Y y Q es operador de X en Z diferenciable en x_0, entonces LQ es diferenciable en x_0 y

$$(LQ)'(x_0) = LQ'(x_0).$$

Observamos, por tanto, que en el cálculo diferencial en espacios de Banach los operadores lineales acotados juegan un papel similar al de las constantes en el cálculo diferencial real o complejo.

Notemos que si F es un operador de $\mathbb{R}^n$ en $\mathbb{R}^m$ que, a una n-tupla $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, le hace corresponder $(F_1(x_1, x_2, \ldots, x_n), F_2(x_1, x_2, \ldots, x_n), \ldots, F_m(x_1, x_2, \ldots, x_n)) \in \mathbb{R}^m$, tenemos que la derivada F' en un punto $x_0 = \left(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}\right)$ se representa por la matriz jacobiana

$$F'(x_0) = \left.\begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} & \cdots & \frac{\partial F_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial x_1} & \frac{\partial F_m}{\partial x_2} & \cdots & \frac{\partial F_m}{\partial x_n} \end{pmatrix}\right|_{x=x_0}.$$

Para ver esto, notamos que la matriz $F'(x_0) = \left(\dfrac{\partial F_i}{\partial x_j}\right)_{x=x_0}$ es un operador lineal acotado de $\mathbb{R}^n$ en $\mathbb{R}^m$, puesto que es una matriz $m \times n$ con coeficientes constantes [38].

A continuación, presentamos ejemplos del cálculo de derivadas de operadores en algunos espacios de Banach.

Ejemplo 1.27. Sea el operador F en $\mathcal{C}([0,1])$ definido por

$$[F(x)](s) = x\int_0^1 h(s,t)x(t)\,dt,$$

donde $h(s,t) = \dfrac{s}{s+t}$ y $0 \leq s \leq 1$. Como

$$\begin{aligned}
\left[F(x_0+hx) - F(x_0)\right](s) &= x_0(s)\int_0^1 \frac{s}{s+t}(hx)(t)\,dt \\
&\quad + (hx)(s)\int_0^1 \frac{s}{s+t}x_0(t)\,dt + (hx)(s)\int_0^1 \frac{s}{s+t}(hx)(t)\,dt, \\
\lim_{h\to 0}\frac{\left[F(x_0+hx) - F(x_0)\right](s)}{h} &= \lim_{h\to 0}\Big(x_0(s)\int_0^1 \frac{s}{s+t}x(t)\,dt \\
&\quad + x(s)\int_0^1 \frac{s}{s+t}x_0(t)\,dt + x(s)\int_0^1 \frac{s}{s+t}(hx)(t)\,dt\Big) \\
&= x_0(s)\int_0^1 \frac{s}{s+t}x(t)\,dt + x(s)\int_0^1 \frac{s}{s+t}x_0(t)\,dt \\
&= L(x)(s)
\end{aligned}$$

y L es un operador lineal acotado en $\mathcal{C}([0,1])$, entonces F es diferenciable Gateaux en x_0. Además, como

$$\lim_{\|v\|\to 0}\frac{\|F(x_0+v) - F(x_0) - Lv\|}{\|v\|} = 0,$$

F es diferenciable Fréchet en x_0.

Ejemplo 1.28. Definimos el operador integral de tipo Hammerstein mixto F en $\mathcal{C}([a,b])$ por

$$[F(x)](s) = x(s) - f(s) - \int_a^b G(s,t)H(t,x(t))\,dt, \quad s\in[a,b], \tag{1.11}$$

para $x\in\mathcal{C}([a,b])$, donde f es una función dada, $G(s,t)$ es el núcleo de un operador integral lineal en $\mathcal{C}([a,b])$ y $H(t,u)$ es una función continua para $t\in[a,b]$ y $u\in(-\infty,\infty)$. Como

$$\begin{aligned}
[F(x_0+hx) - F(x_0)](s) &= hx(s) - \int_a^b G(s,t)\Big(H(t,x_0(t)+hx(t)) - H(t,x_0(t))\Big)\,dt, \\
\lim_{h\to 0}\frac{[F(x_0+hx) - F(x_0)](s)}{h} &= x(s) - \int_a^b G(s,t)\left(\lim_{h\to 0}\frac{H(t,x_0(t)+hx(t)) - H(t,x_0(t))}{h}\right)dt \\
&= x(s) - \int_a^b G(s,t)H_2'(t,x_0(t))\,x(t)\,dt \\
&= [L(x)](s),
\end{aligned}$$

y L es un operador lineal acotado en $\mathcal{C}([a,b])$, entonces F es diferenciable Gateaux en x_0. Además, como

$$\lim_{\|v\|\to 0}\frac{\|F(x_0+v) - F(x_0) - Lv\|}{\|v\|} = 0,$$

F es diferenciable Fréchet en x_0.

Ejemplo 1.29. Consideramos ahora el operador

$$[F(x)](t) = \frac{d^2x(t)}{dt^2} + \phi(x(t)) \tag{1.12}$$

de $\mathcal{C}^2([a,b])$ en $\mathcal{C}([a,b])$, donde $\phi(u)$ es una función continua para $u \in (-\infty, \infty)$. Como

$$\begin{aligned}
[F(x_0+hx)-F(x_0)](t) &= h\,\frac{d^2x(t)}{dt^2} + \phi(x_0(t)+hx(t)) - \phi(x_0(t)),\\
\lim_{h\to 0}\frac{[F(x_0+hx)-F(x_0)](t)}{h} &= \frac{d^2x(t)}{dt^2} + \lim_{h\to 0}\frac{\phi(x_0(t)+hx(t))-\phi(x_0(t))}{h}\\
&= \frac{d^2x(t)}{dt^2} + \phi'(x_0(t))x(t)\\
&= [L(x)](t),
\end{aligned}$$

y L es un operador lineal acotado en $\mathcal{C}([a,b])$, entonces F es diferenciable Gateaux en x_0. Además, como

$$\lim_{\|v\|\to 0}\frac{\|F(x_0+v)-F(x_0)-Lv\|}{\|v\|} = 0,$$

F es diferenciable Fréchet en x_0.

Por otra parte, uno de los resultados que se pierden al pasar de los números reales a espacios más generales es el Teorema del Valor Medio, que asegura que si f es una función diferenciable en un intervalo $[a,b]$, entonces

$$f(b)-f(a) = f'(\xi)(b-a), \quad \text{donde} \quad \xi \in (a,b).$$

Veremos ahora una desigualdad que generaliza al Teorema del Valor Medio. Para ello, dados dos puntos x e y, definimos el segmento que los une como

$$[x,y] = \{x+\theta(y-x); \theta \in [0,1]\}.$$

Diremos que un conjunto Ω es convexo si, para cualesquiera $x, y \in \Omega$, el segmento $[x,y]$ que une x e y está contenido en Ω.

Teorema 1.30. (Teorema del Valor Medio) *Sean X e Y dos espacios de Banach y $F : X \to Y$ un operador diferenciable en un conjunto convexo $\Omega \subset X$. Entonces, si $x_0, x_1 \in \Omega$, se tiene*

$$\|F(x_1)-F(x_0)\| \leq \sup_{0<\theta<1}\|F'(x_0+\theta(x_1-x_0))\|\,\|x_1-x_0\|.$$

Demostración. Dada una función lineal acotada $g : Y \to \mathbb{R}$, definimos:

$$\phi(t) = g(F(x_0+tv)), \quad v = x_1-x_0, \quad t \in [0,1].$$

Teniendo en cuenta que $g' = g$ por ser lineal, se obtiene

$$\phi'(t) = g\left(F'(x_0+tv)(v)\right).$$

Aplicando ahora el Teorema del Valor Medio a ϕ, se sigue

$$\phi(1)-\phi(0) = \phi'(\theta), \quad 0<\theta<1,$$

y sustituyendo por sus expresiones, se tiene

$$g\left(F(x_1)-F(x_0)\right) = g\left(F'(x_0+\theta v)(v)\right).$$

Por tanto,

$$|g\left(F(x_1)-F(x_0)\right)| \leq \|g\| \sup_{0<\theta<1} \|F'\left(x_0+\theta v\right)\| \|v\|.$$

Por el Teorema de Hanh-Banach [29], podemos elegir $g \neq 0$ tal que

$$g\left(F(x_1)-F(x_0)\right) = \|g\| \|F(x_1)-F(x_0)\|.$$

Y finalmente,

$$\|F(x_1)-F(x_0)\| \leq \sup_{0<\theta<1} \|F'\left(x_0+\theta(x_1-x_0)\right)\| \, \|x_1-x_0\|. \quad \blacksquare$$

Corolario 1.31. *Con la notación anterior se tiene que*

$$\|F(x_1)-F(x_0)-F'(\bar{x})(x_1-x_0)\|$$
$$\leq \sup_{0<\theta<1} \|F'\left(x_0+\theta(x_1-x_0)\right)-F'(\bar{x})\| \, \|x_1-x_0\| \quad \textit{con} \quad \bar{x} \in [x_0,x_1].$$

La demostración del corolario anterior se obtiene aplicando el teorema 1.30 al operador $F - F'(\bar{x})$.

Si F es un operador diferenciable entre dos espacios de Banach X e Y, entonces podemos interpretar F' como un operador de X en $\mathcal{L}(X,Y)$. El espacio $\mathcal{L}(X,Y)$ también es un espacio de Banach con la norma

$$\|L\| = \sup_{\|x\|=1} \|Lx\|.$$

En consecuencia, se puede pensar en derivar el operador F'. El resultado de derivar este operador en un punto x_0 es lo que se conoce como la derivada segunda de F en x_0 y se suele denotar por $F''(x_0)$.

Observamos también que $F''(x_0) \in \mathcal{L}\left(X, \mathcal{L}(X,Y)\right)$, y así tenemos que F'' es un operador de X en $\mathcal{L}\left(X, \mathcal{L}(X,Y)\right)$, que también es un espacio de Banach. Si volvemos a derivar este operador en un punto x_0, obtenemos la derivada tercera $F'''(x_0)$. Continuando este proceso, en el caso de que existan las derivadas que van apareciendo, se obtienen las derivadas de órdenes superiores, cuyo estudio no vamos a desarrollar, dado que no se van a utilizar a lo largo de todo este texto. Los estudiantes interesados pueden consultar [27].

1.3. Integración de operadores

A continuación comentamos algunos aspectos sobre el cálculo integral en espacios de Banach. En primer lugar, definimos la integral en el sentido de Riemann para una función de variable real y con valores en un espacio de Banach. A continuación, apoyándonos en la definición anterior, definimos la integral de una función entre dos espacios de Banach en general.

Definición 1.32. *Sea F definida en un intervalo real $[a,b]$ y con valores en un espacio de Banach Y. Entonces podemos definir la integral de F como el límite de la siguiente suma*

$$\sum_{k=0}^{n-1} F(\tau_k)(t_{k+1}-t_k),$$

donde $a = t_0 < t_1 < \cdots < t_n = b$ y $\tau_k \in [t_k, t_{k+1}]$, cuando $\max\limits_{(k)}\{t_{k+1} - t_k\}$ tiende a cero. Si el límite anterior existe, lo llamamos integral de F y lo denotamos por

$$\int_a^b F(t)\, dt.$$

Evidentemente, si la integral existe, es un elemento de Y. Una condición suficiente para que exista dicha integral es que la función F sea continua en el intervalo cerrado $[a, b]$.

Las propiedades de esta integral se deducen de las conocidas para la integral de Riemann en el caso real. De entre ellas, destacamos las siguientes tres por su utilidad [27]:

(*i*) Si consideramos un operador lineal y acotado $L \in \mathcal{L}(Y, Z)$, entonces

$$\int_a^b L(F(t))\, dt = L\left(\int_a^b F(t)\, dt\right).$$

(*ii*) Si $F(t) = \phi(t)y_0$, donde y_0 es un elemento fijo de Y y ϕ es una función real integrable, entonces

$$\int_a^b F(t)\, dt = y_0 \int_a^b \phi(t)\, dt.$$

(*iii*) $\left\|\int_a^b F(t)\, dt\right\| \leq \int_a^b \|F(t)\|\, dt.$

Definición 1.33. *Supongamos ahora que T es un operador definido en un segmento $[x_0, x_1] \subseteq X$ y con valores en el espacio $\mathcal{L}(X, Y)$. En este caso, definimos:*

$$\int_{x_0}^{x_1} T(x)\, dx = \int_0^1 T\left(x_0 + t(x_1 - x_0)\right)(x_1 - x_0)\, dt.$$

Notemos que la función $T\left(x_0 + [\cdot](x_1 - x_0)\right)(x_1 - x_0)$ está en las condiciones de la definición 1.32 con $[a, b] = [0, 1]$.

Si en la definición anterior se considera el caso particular $T = F'$, donde F es un operador de X en Y que tiene derivada continua en el segmento $[x_0, x_1]$, tenemos el siguiente teorema que generaliza la conocida regla de Barrow del cálculo.

Teorema 1.34. *Si F es un operador de X en Y con derivada continua en el segmento $[x_0, x_1] \subset X$, entonces*

$$\int_{x_0}^{x_1} F'(x)\, dx = F(x_1) - F(x_0). \tag{1.13}$$

Demostración. Sea $0 = t_0 < t_1 < \cdots < t_n = 1$ una partición del intervalo $[0, 1]$ con $\lambda = \max\limits_{0 \leq k \leq n-1}\{t_{k+1} - t_k\}$ y $\tau_k \in [t_k, t_{k+1}]$. Entonces,

$$\begin{aligned} \int_{x_0}^{x_1} F'(x)\, dx &= \lim_{\lambda \to 0} \sum_{k=0}^{n-1} F'\left(x_0 + \tau_k(x_1 - x_0)\right)(x_1 - x_0)(t_{k+1} - t_k) \\ &= \lim_{\lambda \to 0} \sum_{k=0}^{n-1} F'(\overline{x}_k)\Delta x_k, \end{aligned}$$

donde $\overline{x}_k = x_0 + \tau_k(x_1 - x_0)$ y $\Delta x_k = (t_{k+1} - t_k)(x_1 - x_0)$ para $k = 0, \ldots, n-1$. Por otro lado,

$$\begin{aligned} F(x_1) - F(x_0) &= \sum_{k=0}^{n-1} \left[F\left(x_0 + t_{k+1}(x_1 - x_0)\right) - F\left(x_0 + t_k(x_1 - x_0)\right)\right] \\ &= \sum_{k=0}^{n-1} \left[F(x_{k+1}) - F(x_k)\right], \end{aligned}$$

donde $x_k = x_0 + t_k(x_1 - x_0)$.

Aplicando el corolario 1.31, como $\overline{x}_k \in [x_k, x_{k+1}]$, se tiene

$$\begin{aligned} \|F(x_{k+1}) - F(x_k) - F'(\overline{x}_k)\Delta x_k\| &\leq \|\Delta x_k\| \sup_{0<\theta<1} \|F'(x_k + \theta\Delta x_k) - F'(\overline{x}_k)\| \\ &= \|x_1 - x_0\|(t_{k+1} - t_k) \sup_{0<\theta<1} \|F'(x_k + \theta\Delta x_k) - F'(\overline{x}_k)\|, \end{aligned}$$

de manera que

$$\begin{aligned} \left\|\sum_{k=0}^{n-1} \left[F(x_{k+1}) - F(x_k) - F'(\overline{x}_k)\Delta x_k\right]\right\| &\leq \sum_{k=0}^{n-1} \|F(x_{k+1}) - F(x_k) - F'(\overline{x}_k)\Delta x_k\| \\ &\leq \|x_1 - x_0\| \sum_{k=0}^{n-1} (t_{k+1} - t_k) \sup_{0<\theta<1} \|F'(x_k + \theta\Delta x_k) - F'(\overline{x}_k)\|. \end{aligned}$$

Como F' es continuo y también uniformemente continuo en $[x_0, x_1]$, [27], considerando el limite cuando $\lambda \to 0$, se obtiene el resultado. ■

Corolario 1.35. (Integración por partes) *Si $P(t)$ es un operador de $[a, b]$ en Y y $Q(t)$ es una función escalar tales que $P'(t)$ y $Q'(t)$ son integrables, entonces*

$$P(b)Q(b) - P(a)Q(a) = \int_a^b P(t)Q'(t)\,dt + \int_a^b P'(t)Q(t)\,dt. \tag{1.14}$$

Demostración. Si $F(t) = P(t)Q(t)$, entonces $F'(t) = P(t)Q'(t) + P'(t)Q(t)$. Aplicando ahora el teorema 1.34, tenemos

$$\int_a^b F'(t)\,dt = F(b) - F(a) = P(b)Q(b) - P(a)Q(a).$$

Por otro lado, de

$$\int_a^b F'(t)\,dt = \int_a^b \left(P(t)Q'(t) + P'(t)Q(t)\right)\,dt = \int_a^b P(t)Q'(t)\,dt + \int_a^b P'(t)Q(t)\,dt,$$

se deduce el resultado. ■

De la propiedad (iii) de la integral se sigue el siguiente resultado que nos permite acotar la integral de un operador que esté acotado por una función real.

Lema 1.36. *Sea T un operador en las condiciones de la definición 1.33 y $\phi(t)$ una función real definida en $[0, 1]$ e integrable. Si*

$$\|T\left(x_0 + \tau(x_1 - x_0)\right)\| \leq \phi\left(t_0 + \tau(t_1 - t_0)\right),\ \tau \in [0, 1], \qquad y \qquad \|x_1 - x_0\| \leq t_1 - t_0,$$

entonces

$$\left\|\int_{x_0}^{x_1} T(x)\,dx\right\| \leq \int_{t_0}^{t_1} \phi(t)\,dt.$$

DEMOSTRACIÓN. En efecto,

$$\begin{aligned}
\left\|\int_{x_0}^{x_1} T(x)\,dx\right\| &= \left\|\int_0^1 T\left(x_0+\tau(x_1-x_0)\right)(x_1-x_0)\,d\tau\right\| \\
&\leq \int_0^1 \|T\left(x_0+\tau(x_1-x_0)\right)\|\,\|x_1-x_0\|\,d\tau \\
&\leq \int_0^1 \phi\left(t_0+\tau(t_1-t_0)\right)(t_1-t_0)\,d\tau \\
&= \int_{t_0}^{t_1} \phi(t)\,dt. \quad \blacksquare
\end{aligned}$$

Como caso particular del lema anterior, tenemos el siguiente resultado.

Corolario 1.37. *Si se cumple la desigualdad*

$$\|T(x)\| \leq \phi(t),$$

para x y t tales que $\|x-x_0\| \leq t-t_0$, entonces

$$\left\|\int_{x_0}^{x_1} T(x)\,dx\right\| \leq \int_{t_0}^{t_1} \phi(t)\,dt,$$

donde x_1 es un elemento tal que $\|x_1-x_0\| \leq t_1-t_0$.

DEMOSTRACIÓN. Si $x \in [x_0,x_1]$ y $t \in [t_0,t_1]$, entonces

$$x = x_0+\tau(x_1-x_0) \quad \text{y} \quad t = t_0+\tau(t_1-t_0) \quad \text{con} \quad \tau \in [0,1],$$

de manera que

$$\|x-x_0\| = \tau\|x_1-x_0\| \leq \tau(t_1-t_0) = t-t_0, \quad \tau \in [0,1].$$

Por consiguiente,

$$\|T\left(x_0+\tau(x_1-x_0)\right)\| \leq \phi\left(t_0+\tau(t_1-t_0)\right), \quad \tau \in [0,1],$$

y se satisfacen las condiciones del lema anterior, del que se sigue el resultado. $\blacksquare$

Terminamos con el Teorema de Taylor. Aunque hay varios enunciados similares [13], hemos elegido el siguiente por su comodidad a la hora de utilizarlo.

Teorema 1.38. (Teorema de Taylor) *Supongamos que F es un operador n veces diferenciable en la bola $B(x_0,r)$ y que $F^{(n)}$ es integrable en el segmento $[x_0,x_1]$ con $x_1 \in B(x_0,r)$. Entonces,*

$$F(x_1) = F(x_0) + \sum_{k=1}^{n-1} \frac{1}{k!}F^{(k)}(x_0)(x_1-x_0)^k + R_n(x_0,x_1), \tag{1.15}$$

donde

$$\begin{aligned}
R_n(x_0,x_1) &= \frac{1}{(n-1)!}\int_{x_0}^{x_1} F^{(n)}(x)(x_1-x)^{n-1}\,dx \\
&= \frac{1}{(n-1)!}\int_0^1 F^{(n)}\left(x_0+t(x_1-x_0)\right)(x_1-x_0)^n(1-t)^{n-1}\,dt.
\end{aligned} \tag{1.16}$$

Demostración. Procedemos por inducción. Para $n = 1$, por el teorema 1.34, se cumple trivialmente. Suponemos que es cierto para $n = m-1$. Consideramos ahora la función escalar

$$Q(t) = \frac{(1-t)^{m-1}}{(m-1)!},$$

cuya derivada es

$$Q'(t) = -\frac{(1-t)^{m-2}}{(m-2)!}. \tag{1.17}$$

Además, si

$$P(t) = F^{(m-1)}\left(x_0 + t(x_1 - x_0)\right)(x_1 - x_0)^{m-1}, \tag{1.18}$$

entonces

$$P'(t) = F^{(m)}\left(x_0 + t(x_1 - x_0)\right)(x_1 - x_0)^{m}.$$

De (1.16), (1.17) y (1.18), se sigue

$$R_{m-1}(x_0, x_1) = -\int_0^1 P(t)Q'(t)\,dt.$$

Aplicando ahora la fórmula de integración por partes (1.14), se deduce

$$\begin{aligned}
R_{m-1}(x_0, x_1) &= -P(1)Q(1) + P(0)Q(0) + \int_0^1 P'(t)Q(t)\,dt \\
&= \frac{1}{(m-1)!}\,F^{(m-1)}(x_0)(x_1 - x_0)^{m-1} \\
&\quad + \frac{1}{(m-1)!}\int_0^1 F^{(m)}\left(x_0 + t(x_1 - x_0)\right)(x_1 - x_0)^{m}(1-t)^{m-1}\,dt \\
&= \frac{1}{(m-1)!}\,F^{(m-1)}(x_0)(x_1 - x_0)^{m-1} + R_m(x_0, x_1).
\end{aligned}$$

Esto establece (1.15) por inducción, con lo que queda probado el teorema. ■

Notemos que (1.15) es equivalente a

$$F(x_1) = F(x_0) + \sum_{k=1}^{n} \frac{1}{k!} F^{(k)}(x_0)(x_1 - x_0)^k + \tilde{R}_n(x_0, x_1),$$

donde

$$\tilde{R}_n(x_0, x_1) = \frac{1}{(n-1)!}\int_{x_0}^{x_1} \left(F^{(n)}(x) - F^{(n)}(x_0)\right)(x_1 - x)^{n-1}\,dx.$$

1.4. Diferencias divididas

En esta sección trataremos resumidamente alguno de los conceptos que posteriormente serán utilizados en el desarrollo de este texto. Somos conscientes de que lo aquí abordado necesita de un desarrollo más detallado del presentado. Es por ello que fundamentalmente nos esforzaremos simplemente en resumir ordenadamente algunos de los resultados que posteriormente se utilizan.

El concepto de diferencia dividida en espacios de Banach fue introducido por J. Schröder [42] generalizando el concepto de diferencia dividida de una función escalar de la misma forma que la derivada Fréchet de un operador en espacios de Banach generaliza el de la derivada de una función escalar.

Así, de forma natural, al considerar una diferencia dividida como una aproximación de una derivada, si F es un operador entre dos espacios de Banach X e Y, a partir de la estimación

$$F(x) \approx F(y) + F'(y)(x-y) \quad \Rightarrow \quad F(x) - F(y) \approx F'(y)(x-y)$$

y teniendo en cuenta que $F'(y) \in \mathcal{L}(X,Y)$, es claro que una diferencia dividida en espacios de Banach, al ser una aproximación de la derivada, debe ser un operador lineal acotado de X en Y que, al aplicarlo a $(x-y)$, es igual a $F(x) - F(y)$. Definimos formalmente a continuación el concepto de diferencia dividida de primer orden de un operador F en espacios de Banach.

Definición 1.39. *Sean X e Y dos espacios de Banach, $F : \Omega \subset X \to Y$ un operador y Ω un subconjunto abierto convexo no vacío de X. Se dice que un operador $\Theta \in \mathcal{L}(X,Y)$ es una diferencia dividida de primer orden de F en los puntos x e y, $x \neq y$, si*

$$\Theta(x-y) = F(x) - F(y). \tag{1.19}$$

En adelante, si $\Theta \in \mathcal{L}(X,Y)$ es una diferencia dividida de primer orden del operador F en los puntos x e y, la denotaremos también por $\Theta = [x, y; F]$.

La condición (1.19) no determina la unicidad de la diferencia dividida a menos que la dimensión de X sea uno (diferencia dividida en $\mathbb{R}$). Se puede probar que si $\dim(X) = d$ y $\dim(Y) = d'$, entonces existen $d'(d-1)+1$ operadores lineales de E_1 en E_2, donde E_1 es un espacio normado completo y E_2 es un espacio normado, cumpliendo (1.19) y que son linealmente independientes [37].

Podemos consultar [7] para la existencia de diferencias divididas de un operador y [46] para ejemplos en algunos espacios particulares. A continuación, damos dos ejemplos en espacios de dimensión finita.

Ejemplo 1.40. Consideramos el caso $X = Y = \mathbb{R}^2$ y denotamos las componentes del operador F por F_1 y F_2. Es decir,

$$\text{para} \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2, \quad \text{tenemos} \quad F(x) = \begin{pmatrix} F_1(x_1,x_2) \\ F_2(x_1,x_2) \end{pmatrix}.$$

Entonces, cada uno de los operadores lineales A_1 y A_2, dados respectivamente por las matrices

$$A_1 = \begin{pmatrix} \dfrac{F_1(x_1,y_2) - F_1(y_1,y_2)}{x_1 - y_1} & \dfrac{F_1(x_1,x_2) - F_1(x_1,y_2)}{x_2 - y_2} \\ \dfrac{F_2(x_1,y_2) - F_2(y_1,y_2)}{x_1 - y_1} & \dfrac{F_2(x_1,x_2) - F_2(x_1,y_2)}{x_2 - y_2} \end{pmatrix}$$

$$A_2 = \begin{pmatrix} \dfrac{F_1(x_1,x_2) - F_1(y_1,x_2)}{x_1 - y_1} & \dfrac{F_1(y_1,x_2) - F_1(y_1,y_2)}{x_2 - y_2} \\ \dfrac{F_2(x_1,x_2) - F_2(y_1,x_2)}{x_1 - y_1} & \dfrac{F_2(y_1,x_2) - F_2(y_1,y_2)}{x_2 - y_2} \end{pmatrix},$$

verifican (1.19).

Por otra parte, si F es diferenciable y su derivada Fréchet F' es continua en el segmento $[x,y]=\{x+t(y-x); t\in[0,1]\}$, el operador lineal dado por

$$A_3=\int_0^1 F'(x+t(y-x))\,dt$$

también verifica (1.19). Esto significa que cualquiera de los tres operadores anteriores son diferencias divididas del operador F en los puntos x e y. Además, cualquier combinación convexa de A_1, A_2 y A_3 es también una diferencia dividida de F en x e y.

Ejemplo 1.41. Sea D un dominio abierto de $\mathbb{R}^n$ y F un operador definido en D con valores en $\mathbb{R}^n$ y tal que $F(x)=(F_1(x),\dots,F_n(x))$.

Sean x e y dos puntos distintos de D y definimos $[x,y;F]$ como la matriz con entradas

$$[x,y;F]_{ij}=\frac{1}{x_j-y_j}\ \left(F_i(x_1,\dots,x_j,y_{j+1},\dots,y_n)-F_i(x_1,\dots,x_{j-1},y_j,\dots,y_n)\right).$$

Es sencillo comprobar que el operador anterior $[x,y;F]\in\mathcal{L}(\mathbb{R}^n,\mathbb{R}^n)$ verifica la condición (1.19). En efecto, vemos que se cumple

$$\left([x,y;F](x-y)\right)_i=\left(F(x)-F(y)\right)_i,$$

puesto que

$$\begin{aligned}
\left([x,y;F](x-y)\right)_i &= \sum_{j=1}^n [x,y;F]_{ij}(x_j-y_j)\\
&= \sum_{j=1}^n \left(F_i(x_1,\dots,x_j,y_{j+1},\dots,y_n)-F_i(x_1,\dots,x_{j-1},y_j,\dots,y_n)\right)\\
&= \left(F_i(x_1,y_2,\dots,y_n)-F_i(y_1,y_2,\dots,y_n)\right)\\
&\quad +\left(F_i(x_1,x_2,y_3,\dots,y_n)-F_i(x_1,y_2,\dots,y_n)\right)\\
&\quad +\cdots+\left(F_i(x_1,\dots,x_{n-1},y_n)-F_i(x_1,\dots,x_{n-2},y_{n-1},y_n)\right)\\
&\quad +\left(F_i(x_1,\dots,x_n)-F_i(x_1,\dots,x_{n-1},y_n)\right)\\
&= F_i(x_1,\dots,x_n)-F_i(y_1,\dots,y_n)\\
&= \left(F(x)-F(y)\right)_i.
\end{aligned}$$

A menudo, se requiere que la aplicación $[\cdot,\cdot;F]:X\times Y\to\mathcal{L}(X,Y)$ tal que $(x,y)\to[x,y;F]\in\mathcal{L}(X,Y)$ satisfaga una condición Lipschitz en algún dominio, lo que implica que el operador F sea diferenciable, tal y como veremos después. Nosotros utilizaremos dicha condición en la forma en que lo hace Laarsonen en [30].

Definición 1.42. *Sean X e Y dos espacios de Banach, $F:\Omega\subset X\to Y$ un operador, Ω un subconjunto abierto convexo no vacío de X y x e y dos puntos distintos de Ω que*

tienen asociada una diferencia dividida de primer orden $[x, y; F] \in \mathcal{L}(X, Y)$. Diremos que la diferencia dividida de primer orden $[x, y; F]$ es Lipschitz continua en Ω si existe una constante $K \geq 0$ tal que

$$\|[x, y; F] - [u, v; F]\| \leq K\left(\|x-u\| + \|y-v\|\right); \quad \forall x, y, u, v \in \Omega; \quad x \neq y, u \neq v. \tag{1.20}$$

Veamos en el siguiente lema que la condición anterior implica que F es diferenciable Fréchet en Ω, [37].

Lema 1.43. *Sean X e Y dos espacios de Banach, $F : \Omega \subset X \to Y$ un operador, Ω un subconjunto abierto convexo no vacío de X y x e y dos puntos distintos de Ω que tienen asociada una diferencia dividida de primer orden $[x, y; F] \in \mathcal{L}(X, Y)$ cumpliendo (1.20). Entonces,*

$$[x, x; F] = F'(x), \tag{1.21}$$

$$\|F'(x) - F'(y)\| \leq 2K\,\|x - y\|, \quad K \geq 0. \tag{1.22}$$

Demostración. Sea $x \in \Omega$. Si $K > 0$, elegimos $\epsilon > 0$ tal que $B(x, \epsilon/K) \subset \Omega$ y denotamos $\delta = \epsilon/K$. Para $\|\Delta x\| \leq \delta$, se tiene entonces

$$\begin{aligned}
\|F(x+\Delta x) - F(x) - [x, x; F]\ (\Delta x)\| &= \|[x+\Delta x, x; F]\ (\Delta x) - [x, x; F](\Delta x)\| \\
&= \|\left([x+\Delta x, x; F] - [x, x; F]\right)\ (\Delta x)\| \\
&\leq \|[x+\Delta x, x; F] - [x, x; F]\|\ \|\Delta x\| \\
&\leq K\ \|\Delta x\|\ \|\Delta x\|.
\end{aligned}$$

Esta condición prueba (1.21) cuando $K \neq 0$ y $\|\Delta x\| \to 0$, puesto que

$$\lim_{\|\Delta x\| \to 0} \frac{\|F(x+\Delta x) - F(x) - [x, x; F](\Delta x)\|}{\|\Delta x\|} = 0.$$

Si $K = 0$, por (1.20), existe un operador $L \in \mathcal{L}(X, Y)$ tal que $[x, y; F] = L$ para todo $x, y \in \Omega$. Entonces, por (1.19), podemos tomar cualquier δ en lo anterior y se tiene $F'(x) = L$.

Para completar la demostración vemos (1.22). Sean $x, y \in \Omega$ y, por (1.20), tenemos

$$\|F'(x) - F'(y)\| = \|[x, x; F] - [y, y; F]\| \leq K\ \left(\|x-y\| + \|x-y\|\right) = 2K\|x-y\|. \quad \blacksquare$$

Notemos que en las condiciones indicadas se tiene que F' cumple una condición Lipschitz de la forma (1.22) con constante de Lipschitz $2K$.

Una consecuencia inmediata de (1.20) y (1.21) es:

$$\|[x, y; F] - F'(z)\| \leq K\ \left(\|x-z\| + \|y-z\|\right), \quad K \geq 0, \quad \forall x, y, z \in \Omega. \tag{1.23}$$

Recíprocamente, si suponemos que F es diferenciable Fréchet en Ω y que su derivada Fréchet satisface (1.22), entonces se sigue que F tiene tiene una diferencia dividida de primer orden Lipschtz contínua en Ω. En efecto, para ello, podemos tomar por ejemplo

$$[x, y; F] = \int_0^1 F'(x + t(y-x))\,dt. \tag{1.24}$$

Sin embargo, con la excepción del caso $\dim(X) = 1$, sabemos que (1.24) no es la única diferencia dividida de primer orden Lipschitz contínua de F.

Ejemplo 1.44. En el ejemplo 1.28 hemos visto que el operador (1.11) es diferenciable con derivada

$$[F'(x_0)\,x](s) = x(s) - \int_a^b G(s,t)\,H_2'(t,x_0(t))\,x(t)\,dt.$$

Si utilizamos (1.24), se sigue

$$([x,y;F]-[u,v;F])\,z(s) = \int_0^1 (F'(y+\tau(x-y)) - F'(v+\tau(u-v)))\,z(s)\,d\tau$$

$$= -\int_0^1 \left(\int_a^b G(s,t)\,(H_2'(t,(y+\tau(x-y))(t)) - H_2'(t,(v+\tau(u-v))(t)))\,z(t)\,dt \right) d\tau.$$

Suponiendo ahora que H_2' es Lipschitz continua con constante de Lipschitz C y teniendo en cuenta

$$\|[x,y;F]-[u,v;F]\| = \sup_{\|z\|=1} \|([x,y;F]-[u,v;F])\ z\,\|\,, \tag{1.25}$$

$$\begin{aligned}
\|([x,y;F]-[u,v;F])\ z\,\| &\leq C \left\| \int_a^b G(s,t)\,dt \right\| \int_0^1 \|y+\tau(x-y)-(v+\tau(u-v))\|\,d\tau\,\|z\| \\
&\leq C \left\| \int_a^b G(s,t)\,dt \right\| \int_0^1 \|\tau(x-u)+(1-\tau)(y-v))\|\,dt\,\|z\| \\
&\leq \frac{C}{2} \left\| \int_a^b G(s,t)\,dt \right\| (\|x-u\|+\|y-v\|)\ \|z\|,
\end{aligned}$$

se sigue que la diferencia dividida de primer orden $[x,y;F]$ es Lipschitz continua en $\mathcal{C}([a,b])$ con constante de Lipschitz $\dfrac{C}{2}\left\| \int_a^b G(s,t)\,dt \right\|$.

En el siguiente teorema se da una caracterización de las diferencias divididas de primer orden de la forma (1.24), [37].

Teorema 1.45. *Sea $[\cdot,\cdot;F] : \Omega\times\Omega \to \mathcal{L}(X,Y)$ un operador que satisface las condiciones (1.19) y (1.20). Las siguientes afirmaciones son equivalentes*

(i) La igualdad (1.24) se cumple para todo $x,y \in \Omega$.

(ii) Para todo par de puntos $u,v \in \Omega$ tales que $2v-u \in \Omega$, se tiene

$$[u,v;F] = 2[u,2v-u;F] - [v,2v-u;F]. \tag{1.26}$$

Demostración. Sustituyendo $w = v-u$, podemos reescribir (1.26) de la forma

$$[u,u+w;F] + [u+w,u+2w;F] = 2\,[u,u+2w;F]\,. \tag{1.27}$$

Entonces, la implicación (i) $\Rightarrow$(ii) se sigue de forma inmediata observando que

$$\begin{aligned}
\int_0^1 F'(u+tw)dt + \int_0^1 F'(u+w+tw)dt &= \lim_n \frac{1}{n}\left[\sum_{j=1}^n F'(u+\frac{j}{n}w) + \sum_{j=1}^n F'(u+w+\frac{j}{n}w) \right] \\
&= 2\lim_n \frac{1}{2n}\sum_{j=1}^{2n} F'(u+\frac{j}{2n}2w) \\
&= 2\int_0^1 F'(u+t2w)dt.
\end{aligned}$$

Para probar la implicación (ii)⇒(i), observamos en primer lugar que (1.27) implica

$$2^n[u, u + 2^n w; F] = \sum_{j=1}^{2^n} [u + (j + 1)w, u + jw; F], \quad \text{para todo } n \in \mathbb{N}. \tag{1.28}$$

Esto se puede probar fácilmente por inducción. Para verlo, consideramos los primeros cuatro sumandos. Teniendo en cuenta

$$[u, u + w; F] + [u + w, u + 2u; F] + [u + 2w, u + 3w; F] + [u + 3w, u + 4w; F]$$

$$= 2\,[u, u + 2w; F] + 2\,[u + 2w, u + 4w; F] = 4\,[u, u + 4w; F]\,,$$

es obvio que se sigue la inducción.

Consideramos ahora la igualdad (1.28) para $u = x$ y $w = 2^{-n}(y - x)$ y utilizando (1.23), obtenemos

$$\begin{aligned} \|\,[x, y; F] - 2^{-n} \textstyle\sum_{k=j}^{2^n} F'(x + jw)\| &= \tfrac{1}{2^n} \|\textstyle\sum_{k=j}^{2^n} ([x + (j - 1)w, x + kw; F] - F'(x + jw))\| \\ &\leq 2^{-n}2^n k\|w\| = \tfrac{1}{2^n} k\|y - x\|. \end{aligned}$$

Ahora, si $n \to \infty$, obtenemos (1.24). ■

A continuación, exponemos brevemente algunas condiciones de tipo Lipschitz para las diferencias divididas de primer orden que han sido utilizadas por otros autores.

La condición Lipschitz (1.20) ha sido utilizada, por ejemplo en [26], [30] y [36]. Utilizando la noción anterior de diferencia dividida, Schmidt en [41] y Sergeev en [43] generalizan el conocido método de la secante a espacios de Banach. Para probar la convergencia de dicho método ambos consideran una condición Lipschitz de la forma:

$$\|[x, y; F] - [y, z; F]\| \leq K\|x - z\|, \quad x, y, z \in \Omega, \quad K \geq 0. \tag{1.29}$$

Es fácil ver que (1.29) implica (1.20), puesto que

$$\|\,[x, y; F] - [u, v; F]\,\| \leq \|\,[x, y; F] - [y, u; F]\,\| + \|\,[y, u; F] - [u, v; F]\,\| \leq K\|x - u\| + K\|y - v\|.$$

Por otro lado, tomando $x = z$ en (1.29), se sigue $[x, y; F] = [y, x; F]$.

Numerosos ejemplos importantes de diferencias divididas satisfacen la condición (1.20), pero no cumplen la relación simétrica anterior y, en consecuencia, no verifican (1.29). Notemos que en el ejemplo 1.40, A_1 y A_2 no son simétricas, mientras que A_3 y $\frac{1}{2}A_1 + \frac{1}{2}A_2$ si que lo son.

Por otra parte, Schmidt [41] y Dennis [18] prueban la convergencia del método de la secante bajo la condición

$$\|[x, y; F] - [y, z; F]\| \leq a\|x - z\| + b\,(\|x - y\| + \|y - z\|)\,, \quad a, b \geq 0. \tag{1.30}$$

Además, sustituyendo la condición (1.20) por (1.30) en el lema 1.43, se obtiene la misma conclusión con $K = a + b$, [18], [34].

Para operadores F cuya derivada F' satisface una condición Lipschitz con constante C, se sigue fácilmente la siguiente estimación para la distancia del operador a su parte local lineal

$$\|F(y) - F(x) - F'(x)(y - x)\| \leq \frac{C}{2}\|y - x\|^2, \tag{1.31}$$

sin más que tener en cuenta

$$F(x) - F(y) - F'(u)(x-y) = \int_0^1 (F'(y+t(x-y)) - F'(u))\, dt(x-y). \tag{1.32}$$

Notemos que si utilizamos una diferencia dividida de primer orden Lipschitz continua con constante K, entonces F' es Lipschitz continua con constante $2K$ (lema 1.43) y (1.31) tiene que ser reemplazada por una de las estimaciones que aparecen en el siguiente resultado.

Teorema 1.46. *Si $[x, y; F]$ es una diferencia dividida de primer orden que satisface (1.23), entonces cada una de las siguientes cuatro expresiones*

$$e_1 = K(\|x-u\| + \|y-u\| + \|u-v\|)\,\|x-y\|,$$

$$e_2 = K(\|x-v\| + \|y-v\| + \|u-v\|)\,\|x-y\|,$$

$$e_3 = K(\|x-y\| + \|y-u\| + \|y-v\|)\,\|x-y\|,$$

$$e_4 = K(\|x-y\| + \|x-u\| + \|x-v\|)\,\|x-y\|,$$

es una estimación para

$$\|F(x) - F(y) - [u, v; F](x-y)\|.$$

Demostración. Será suficiente dar la demostración para e_1 y e_3 porque la demostración de e_2 se obtiene de e_1 intercambiando u y v y, de forma similar, la demostración de e_4 se puede obtener considerando x en lugar de y en los lugares apropiados. Para probar e_1, obsérvese que de (1.32) y

$$\|F'(y+t(x-y)) - F'(u)\| \le 2K(t\|x-u\| + (1-t)\|y-u\|)$$

se sigue

$$\begin{aligned} \|F(x) - F(y) - F'(u)(x-y)\| &\le 2K\left(\tfrac{1}{2}\|x-u\| + \tfrac{1}{2}\|y-u\|\right)\|x-y\| \\ &= K(\|x-u\| + \|y-u\|)\,\|x-y\|. \end{aligned}$$

Además, como

$$\|F'(u) - [u, v; F]\| \le K\|u-v\|,$$

tenemos

$$\|F(x) - F(y) - [u, v; F](x-y)\| \le K(\|x-u\| + \|y-u\| + \|u-v\|)\,\|x-y\|$$

y se cumple e_1.

Para la demostración de e_3 consideramos

$$\begin{aligned} \|F(x) - F(y) - [u, v; F](x-y)\| &\le \|F(x) - F(y) - F'(y)(x-y)\| \\ &\quad + \|(F'(y) - [u, v; F](u, v))(x-y)\| \\ &\le K\|x-y\|^2 + K(\|y-u\| + \|y-v\|)\,\|x-y\|. \quad \blacksquare \end{aligned}$$

A continuación, generalizamos la condición de Lipschitz continuidad para la diferencia dividida $[x, y; F] \in \mathcal{L}(X, Y)$ a una condición de Hölder continuidad, que también implica la diferenciabilidad del operador F.

Definición 1.47. *Sean X e Y dos espacios de Banach, $F:\Omega\subset X\to Y$, Ω un subconjunto abierto convexo no vacío de X y x e y dos puntos distintos de Ω que tienen asociada una diferencia dividida de primer orden $[x,y;F]$ es (K,p)-Hölder continua en Ω si existen constantes $K\geq 0$ y $p\in[0,1]$ tales que*

$$\|[x,y;F]-[u,v;F]\|\leq K\,(\|x-u\|^p+\|y-v\|^p)\,;\ \forall x,y,u,v\in\Omega;\ x\neq y,\ u\neq v. \tag{1.33}$$

Observamos que si $p=1$, entonces $[x,y;F]$ es Lipschitz continua en Ω.

Ejemplo 1.48. En el ejemplo 1.29 hemos visto que el operador (1.12) es diferenciable con derivada

$$[F'(x_0)\,x](t)=\frac{d^2x(t)}{dt^2}+\phi'(x_0(t))\,x(t).$$

Si utilizamos (1.24), se sigue

$$\begin{aligned}([x,y;F]-[u,v;F])\ z(t) &= \int_0^1 \left(F'(y+\tau(x-y))-F'(v+\tau(u-v))\,\right) z(t)\,d\tau\\ &= \int_0^1 \Big(\phi'\Big((y+\tau(x-y))(t)\Big)-\phi'\Big((v+\tau(u-v))(t)\Big)\Big)\ z(t)\,d\tau.\end{aligned}$$

Suponiendo ahora que ϕ' es (C,p)-Hölder continua y teniendo en cuenta (1.25) y

$$\begin{aligned}\|([x,y;F]-[u,v;F])\ z\| &\leq C\int_0^1 \|y+\tau(x-y)-(v+\tau(u-v))\|^p\,d\tau\,\|z\|\\ &= C\int_0^1 \|\tau(x-u)+(1-\tau)(y-v)\|^p\,dt\,\|z\|\\ &\leq \frac{C}{1+p}\,(\|x-u\|^p+\|y-v\|^p)\,\|z\|,\end{aligned}$$

se sigue que la diferencia dividida de primer orden es (K,p)-Hölder continua en $\mathcal{C}([a,b])$ con $K=\frac{C}{1+p}$.

Además, una consecuencia inmediata de (1.21) y (1.33) es:

$$\|[x,y;F]-F'(z)\|\leq K\,(\|x-z\|^p+\|y-z\|^p),\ K\geq 0,\ p\in[0,1],\ \forall x,y,z\in\Omega. \tag{1.34}$$

En el caso (1.34) es conocido que la derivada Fréchet F' existe en Ω, satisface (1.21) y F' es $(2K,p)$-Hölder continua en Ω, [4].

Una generalización interesante de (1.30) es la dada por Argyros en [4] y que se puede resumir en el siguiente resultado.

Lema 1.49. *Sea Ω un conjunto abierto convexo no vacío de X. Suponemos que, para todo $x,y\in\Omega$, existe una diferencia dividida de primer orden $[x,y;F]\in\mathcal{L}(X,Y)$. Si $z\in\Omega$, entonces*

$$\|[x,y;F]-[y,z;F]\|\leq a\ \|x-z\|^p+b\ (\|x-y\|^p+\|y-z\|^p), \tag{1.35}$$

donde $p\in(0,1]$ y $a,b\geq 0$. Además,

(*i*) $[x,x;F]=F'(x)$, $x\in\operatorname{int}(\Omega)$,

(*ii*) $\|F'(x) - F'(y)\| \leq 2(a+b)\|x-y\|^p$.

Demostración. Veamos el apartado (i). Elegimos $x \in \operatorname{int}(\Omega)$ y $\delta > 0$ tales que $B(x,\delta) \subset \Omega$. Para $\|\Delta x\| \leq \delta$ se tiene

$$\begin{aligned}
\|F(x+\Delta x) - F(x) - [x,x;F]\ (\Delta x)\| &= \|[x+\Delta x, x; F]\ (\Delta x) - [x,x;F](\Delta x)\| \\
&= \|\left([x+\Delta x, x; F] - [x,x;F]\right)\ (\Delta x)\| \\
&\leq \|[x+\Delta x, x; F] - [x,x;F]\|\ \|\Delta x\| \\
&\leq (a+b)\ \|(\Delta x)\|^p\ \|\Delta x\|.
\end{aligned}$$

Esta desigualdad prueba (i) cuando $a+b \neq 0$ y $\|\Delta x\| \to 0$, ya que

$$\lim_{\|\Delta x\| \to 0} \frac{\|F(x+\Delta x) - F(x) - [x,x;F](\Delta x)\|}{\|\Delta x\|} = 0.$$

Para solventar el caso $a+b=0$, notamos que $a=b=0$ y por (1.35) existe un $L \in \mathcal{L}(X,Y)$ tal que $[x,y;F] = L$ para cada $x,y \in \Omega$. Por tanto, por (1.19), podemos elegir δ arbitrario y $F'(x) = L$.

El apartado (ii) se sigue de (1.35) y

$$\begin{aligned}
\|F'(x) - F'(y)\| &\leq \|[x,x;F] - [x,y;F]\| + \|[x,y;F] - [y,y;F]\| \\
&\leq a\ \|x-y\|^p + b\ \|x-y\|^p + a\ \|x-y\|^p + b\ \|x-y\|^p \\
&= 2(a+b)\ \|x-y\|^p,
\end{aligned}$$

lo que completa la demostración. ■

Todavía podemos generalizar más las condiciones de Lipschitz y Hölder continuidad para la diferencia dividida $[x,y;F] \in \mathcal{L}(X,Y)$ suavizando ambas condiciones mediante la condición:

$$\|[x,y;F] - [u,v;F]\| \leq \omega(\|x-u\|, \|y-v\|); \quad x,y,v,w \in \Omega, \tag{1.36}$$

donde $\omega : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ es una función continua y no decreciente en sus dos componentes.

En el siguiente teorema veremos que la condición (1.36) implica (1.21) cuando $\omega(0,0) = 0$, [23].

Teorema 1.50. *Sea Ω un conjunto abierto convexo no vacío de un espacio de Banach X. Suponemos que, para cada par de puntos $x,y \in \Omega$, existe una diferencia dividida de primer orden $[x,y;F] \in \mathcal{L}(X,Y)$ satisfaciendo (1.36) y $\omega(0,0) = 0$. Entonces, se verifica la condición (1.21).*

Demostración. Sea $\{x_n\} \subset \Omega$ tal que $\lim\limits_{n\to\infty} x_n = x$. Si consideramos $A_n = [x_n, x; F] \in \mathcal{L}(X,Y)$, entonces

$$\|A_n - A_m\| = \|[x_n, x; F] - [x_m, x; F]\| \leq \omega(\|x_n - x_m\|, 0).$$

Como $\{x_n\}$ es convergente, se tiene que $\{A_n\}$ es una sucesión de Cauchy y, por tanto, existe $\lim\limits_{n\to\infty} A_n = \tilde{A} \in \mathcal{L}(X,Y)$ y podemos definir $[x,x;F] = \tilde{A} = \lim\limits_{n\to\infty} A_n$. Demostramos entonces

que $\tilde{A} = F'(x)$:

$$\begin{aligned}\|F(x+\Delta x) - F(x) - [x,x;F]\ (\Delta x)\| &= \|[x+\Delta x, x; F]\ (\Delta x) - [x,x;F](\Delta x)\| \\ &= \|\left([x+\Delta x, x; F] - [x,x;F]\right)\ (\Delta x)\| \\ &\leq \|[x+\Delta x, x; F] - [x,x;F]\|\ \|\Delta x\| \\ &\leq \omega(\|\Delta x\|, 0)\|\Delta x\|.\end{aligned}$$

En consecuencia,

$$\lim_{\|\Delta x\|\to 0} \frac{\|F(x+\Delta x) - F(x) - [x,x;F](\Delta x)\|}{\|\Delta x\|} \leq \lim_{\|\Delta x\|\to 0} \omega(\|\Delta x\|, 0) = \omega(0,0) = 0. \quad \blacksquare$$

Es fácil ver ahora que la condición (1.36) generaliza la condición (1.33), sin más que considerar $\omega(s,t) = K(s^p + t^p)$.

Por otra parte, también podemos observar que si F no es diferenciable, entonces la función ω dada en (1.36) es tal que $\omega(0,0) > 0$, como puede verse en el siguiente ejemplo.

Ejemplo 1.51. Dado el siguiente sistema de ecuaciones no lineales

$$\begin{cases} x^2 - y + 1 + \frac{1}{9}|x-1| = 0, \\ y^2 + x - 7 + \frac{1}{9}|y| = 0, \end{cases}$$

lo podemos escribir como $F(x) = 0$, donde $x = (x_1, x_2) \in \mathbb{R}^2$ y $F : \mathbb{R}^2 \to \mathbb{R}^2$ con $F = (F_1, F_2)$ y

$$F_1(x_1,x_2) = {x_1}^2 - x_2 + 1 + \frac{1}{9}|x_1 - 1|, \qquad F_2(x_1,x_2) = {x_2}^2 + x_1 - 7 + \frac{1}{9}|x_2|.$$

Tomando $x = (x_1,x_2) \in \mathbb{R}^2$ y la norma $\|x\| = \|x\|_\infty = \max_{1\leq i\leq 2} |x_i|$, la norma correspondiente para $A \in \mathbb{R}^2 \times \mathbb{R}^2$ es

$$\|A\| = \max_{1\leq i\leq 2} \sum_{j=1}^{2} |a_{ij}|.$$

Ahora, si $v, w \in \mathbb{R}^2$, podemos definir $[v,w;F] \in \mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ como, [37],

$$[v,w;F]_{i1} = \frac{F_i(v_1,w_2) - F_i(w_1,w_2)}{v_1 - w_1}, \qquad [v,w;F]_{i2} = \frac{F_i(v_1,v_2) - F_i(v_1,w_2)}{v_2 - w_2}, \qquad i = 1,2.$$

Entonces,

$$[v,w;F] = \begin{pmatrix} \dfrac{v_1^2 - w_1^2}{v_1 - w_1} & -1 \\ 1 & \dfrac{v_2^2 - w_2^2}{v_2 - w_2} \end{pmatrix} + \frac{1}{9}\begin{pmatrix} \dfrac{|v_1 - 1| - |w_1 - 1|}{v_1 - w_1} & 0 \\ 0 & \dfrac{|v_2| - |w_2|}{v_2 - w_2} \end{pmatrix}$$

$$\|[x,y;F] - [v,w;F]\| \leq \|x - v\| + \|y - w\| + \frac{2}{9},$$

de manera que $\omega(s,t) = s + t + \dfrac{2}{9}$ y $\omega(0,0) > 0$. Observamos que esta situación surge del hecho de que la función F sea no diferenciable.

1.5. Métodos iterativos en espacios de Banach

Dado un operador $F : \Omega \subset X \to Y$ definido en un dominio abierto convexo no vacío Ω contenido en un espacio de Banach X y con valores en un espacio de Banach Y, nos planteamos la aproximación de una solución x^* de la ecuación

$$F(x) = 0. \tag{1.37}$$

En estas condiciones tan generales, la ecuación (1.37) puede representar una ecuación escalar, un sistema de ecuaciones, una ecuación diferencial, una ecuación integral, etc.

Encontrar una solución exacta de (1.37) suele ser difícil y, por eso, se recurre habitualmente a su aproximación mediante métodos iterativos de la forma

$$\begin{cases} \text{dados } x_{-k}, x_{-k+1}, \dots, x_{-1}, x_0 \text{ en } \Omega, \\ x_{n+1} = G(x_{n-k}, x_{n-k+1}, \dots, x_{n-1}, x_n), \quad n \geq 0, \end{cases} \tag{1.38}$$

donde $x_{-k}, x_{-k+1}, \dots, x_{-1}, x_0$ son aproximaciones iniciales de x^*. Los métodos iterativos (1.38) con $k = 0$ se llaman métodos iterativos punto a punto (sin memoria) y si $k \geq 1$, se llaman métodos iterativos con memoria.

Un método iterativo de la forma (1.38) persigue que, a partir de las aproximaciones iniciales $x_{-k}, x_{-k+1}, \dots, x_{-1}, x_0$, la sucesión de valores que define, $\{x_n\}$, sea tal que los valores x_n sean mejores aproximaciones de x^* a medida que n crece y que además $\lim_n x_n = x^*$. Esto es lo que llamamos convergencia del método iterativo.

A la hora de estudiar la convergencia de un método iterativo, existen tres tipos de resultados de convergencia: locales, semilocales y globales. En primer lugar, los resultados de convergencia local, que, a partir de determinadas condiciones que tiene que cumplir el operador F y exigiendo condiciones a la solución x^*, proporcionan la denominada bola de convergencia del método iterativo, que nos indica la accesibilidad de x^* a partir de aproximaciones iniciales consideradas en dicha bola. En segundo lugar, tenemos los resultados de convergencia semilocal, que, a partir de determinadas condiciones que tiene que cumplir el operador F y exigiendo condiciones a las aproximaciones iniciales $x_{-k}, x_{-k+1}, \dots, x_{-1}, x_0$, proporcionan el dominio de parámetros asociado a las condiciones que deben satisfacer las aproximaciones iniciales para tener asegurada la convergencia de la sucesión $\{x_n\}$ a x^*. Por último, tenemos los resultados de convergencia global, que, a partir de determinadas condiciones que tiene que cumplir el operador F, e independientemente de las aproximaciones iniciales, aseguran la convergencia de la sucesión $\{x_n\}$ a x^*.

Como vemos, todos los resultados anteriores pasan por exigir determinadas condiciones al operador F. Sin embargo, la exigencia de condiciones para la solución, para las aproximaciones iniciales, o para ninguna de éstas, determinan los diferentes tipos de resultados de convergencia. Por un lado, el estudio de la convergencia local tiene el inconveniente de tener que asegurar que la solución x^*, que desconocemos, satisfaga determinadas condiciones. Por otro lado, la ausencia de condiciones para las aproximaciones iniciales, e incluso para la solución x^*, hace que el estudio de la convergencia global sea, en general, excesivamente particular en cuanto al tipo de operadores a tratar.

Por otra parte, cuando estudiamos la aplicabilidad de un método iterativo (1.38) para resolver (1.37), aparece un problema importante: la localización de aproximaciones iniciales $x_{-k}, x_{-k+1}, \dots, x_{-1}, x_0$ suficientemente buenas como para que la sucesión $\{x_n\}$, dada por

(1.38), converja a una solución empezando en ellas. El conjunto de estas aproximaciones iniciales es lo que permite medir la accesibilidad del método iterativo (1.38). Podemos observar la accesibilidad del método (1.38) de forma experimental, mediante cuencas de atracción y regiones de accesibilidad, al considerar ecuaciones particulares, y/o de forma teórica, mediante dominios de parámetros, para cualquier ecuación.

La cuenca de atracción es el conjunto de todos los puntos de salida a partir de los cuales el método iterativo (1.38) converge a una solución de una ecuación particular una vez fijada una tolerancia o un número máximo de iteraciones. La región de accesibilidad permite visualizar, a partir de condiciones de convergencia, qué puntos podemos utilizar como puntos de salida para que el método iterativo converja, empezando en ellos, a una solución de una ecuación particular. Y el dominio de parámetros establece gráficamente, en el plano real, la relación entre los parámetros iniciales que se definen a partir de las condiciones impuestas a los puntos de salida en un resultado teórico de convergencia.

En este texto vamos a hacer hincapié solo en el estudio de la convergencia semilocal de los métodos iterativos que aparecen en él, pero realizándolo mediante diferentes técnicas de demostración y viendo qué es lo que aporta cada una de ellas en el estudio de los métodos iterativos aquí presentados.

También sabemos que las propiedades de convergencia dependen de la elección de la distancia $\|\cdot\|$, pero la velocidad de convergencia de una sucesión $\{x_n\}$, para una distancia dada, se caracteriza por la velocidad de convergencia de la sucesión de números no negativos $\{\|x_n - x^*\|\}$, donde $x^* = \lim\limits_{n\to+\infty} x_n$. Una medida muy utilizada habitualmente de la velocidad de convergencia es el R-orden de convergencia, que definimos a continuación.

Definición 1.52. *Sean $\{x_n\}$ una sucesión en un espacio de Banach X que converge a un punto $x^* \in X$, un número $\tau \geq 1$ y*

$$e_n(\tau) = \begin{cases} n & si\ \tau = 1, \\ \tau^n & si\ \tau > 1, \end{cases} \qquad n \geq 0.$$

i) Decimos que τ es un R-orden de convergencia de la sucesión $\{x_n\}$ si existen dos constantes $b \in (0,1)$ y $B \in (0,+\infty)$ tales que

$$\|x_n - x^*\| \leq Bb^{e_n(\tau)}.$$

ii) Decimos que τ es el R-orden de convergencia de la sucesión $\{x_n\}$ si existen constantes $a, b \in (0,1)$ y $A, B \in (0,+\infty)$ tales que

$$Aa^{e_n(\tau)} \leq \|x_n - x^*\| \leq Bb^{e_n(\tau)}, \quad n \geq 0.$$

En general, comprobar la doble acotación de *ii*) es complicado, por eso normalmente solo se buscan acotaciones superiores como las de *i*). Por tanto, si encontramos un R-orden de convergencia τ de la sucesión $\{x_n\}$, entonces se dice que la sucesión $\{x_n\}$ tiene R-orden de convergencia al menos τ. Este es el argumento utilizado habitualmente para estudiar el orden de convergencia de un método iterativo en espacios de Banach.

A continuación, vamos a focalizar todo lo que acabamos de describir para un método iterativo concreto, el método de Newton, con el objetivo de mejorar su entendimiento.

1.5.1. Convergencia semilocal del método de Newton

De entre los métodos iterativos de la forma (1.38), que se pueden utilizar para aproximar una solución x^* de la ecuación (1.37), el más conocido de todos ellos es, sin lugar a dudas, el método de Newton, que destaca por su sencillez, fácil aplicación y eficiencia. Este método tiene el siguiente algoritmo:

$$\begin{cases} \text{dado } x_0 \text{ en } \Omega, \\ x_{n+1} = x_n - [F'(x_n)]^{-1}F(x_n), \quad n \geq 0. \end{cases}$$

El primer resultado de convergencia semilocal para el método de Newton en espacios de Banach fue dado por Kantorovich [27], se conoce universalmente como teorema de Newton-Kantorovich y se demuestra, bajo las condiciones que se exigen a continuación al operador F y al punto inicial x_0, mediante el conocido principio de la mayorante.

En primer lugar, suponemos que se satisfacen las siguientes tres condiciones:

(A1) $\|F(x_0)\| \leq \delta$,

(A2) existe el operador $[F'(x_0)]^{-1} \in \mathcal{L}(X,Y)$, para algún $x_0 \in \Omega$, y es tal que $\|[F'(x_0)]^{-1}\| \leq \theta$,

(A3) $\|F'(x) - F'(y)\| \leq C\|x-y\|,\ C \geq 0,\ x,y \in \Omega$.

En segundo lugar, definimos el polinomio cuadrático

$$p(t) = \frac{C}{2}t^2 - \frac{t}{\theta} + \delta, \qquad t \in [0,t'],$$

tal que $t^* \leq t^{**} < t'$, donde $t^* = \frac{1-\sqrt{1-2C\delta\theta^2}}{C\theta}$ es la raíz positiva más pequeña del polinomio p y $t^{**} = \frac{1+\sqrt{1-2C\delta\theta^2}}{C\theta}$ es la raíz positiva más grande, y la sucesión escalar $\{t_n\}$ dada por

$$\begin{cases} t_0 = 0, \\ t_{n+1} = t_n - \dfrac{p(t_n)}{p'(t_n)}, \quad n \geq 0. \end{cases}$$

Observemos que la sucesión $\{t_n\}$ converge de forma creciente a t^*, sin más que tener en cuenta que es una sucesión monótona no decreciente y está acotada superiormente por t^*.

Teorema 1.53. (Teorema de Newton-Kantorovich) *Sea $F : \Omega \subset X \to Y$ un operador una vez continuamente diferenciable Fréchet, definido en un dominio abierto convexo no vacío Ω de un espacio de Banach X y con valores en un espacio de Banach Y. Supongamos que se verifican las condiciones (A1)–(A3), $C\delta\theta^2 \leq \frac{1}{2}$ y $B(x_0,t^*) \subset \Omega$, donde $t^* = \frac{1-\sqrt{1-2C\delta\theta^2}}{C\theta}$. Entonces, la ecuación $F(x) = 0$ tiene una solución x^* y el método de Newton converge a x^* empezando en x_0. Además, $x_n, x^* \in \overline{B(x_0,t^*)}$, para todo $n \in \mathbb{N}$. Por otra parte, si $C\delta\theta^2 < \frac{1}{2}$, x^* es la única solución de $F(x) = 0$ en $B(x_0,t^{**}) \cap \Omega$, donde $t^{**} = \frac{1+\sqrt{1-2C\delta\theta^2}}{C\theta}$.*

Demostración. En primer lugar, x_1 está bien definido, puesto que, por hipótesis, existe $[F'(x_0)]^{-1}$. Además,

$$\|x_1 - x_0\| = \|[F'(x_0)]^{-1}F(x_0)\| \leq \|[F'(x_0)]^{-1}\|\|F(x_0)\| \leq -\frac{p(t_0)}{p'(t_0)} = t_1 - t_0 < t^*$$

y, en consecuencia, $x_1 \in B(x_0, t^*)$.

A continuación, veamos que x_2 está bien definido. Para ello, tenemos

$$\begin{aligned} \|I - [F'(x_0)]^{-1}F'(x_1)\| &\leq \|[F'(x_0)]^{-1}\|\|F'(x_0) - F'(x_1)\| \\ &\leq C\|[F'(x_0)]^{-1}\|\|x_1 - x_0\| \\ &\leq -\frac{C}{p'(t_0)}(t_1 - t_0) \\ &= -\frac{1}{p'(t_0)}(p'(t_1) - p'(t_0)) \\ &= 1 - \frac{p'(t_1)}{p'(t_0)} \\ &< 1, \end{aligned}$$

por ser $p'(t_1) > 0$ y $p'(t_1) > p'(t_0)$. Luego, por el lema de Banach (lema 1.22), existe $[F'(x_1)]^{-1}$ y

$$\|[F'(x_1)]^{-1}\| \leq -\frac{1}{p'(t_1)}.$$

Por tanto, x_2 está bien definido.

Por otra parte, como

$$\begin{aligned} F(x_1) &= F(x_0) + F'(x_0)(x_1 - x_0) + \int_{x_0}^{x_1} (F'(z) - F'(x_0))\, dz \\ &= \int_{x_0}^{x_1} (F'(z) - F'(x_0))\, dz \\ &= \int_0^1 (F'(x_0 + \tau(x_1 - x_0)) - F'(x_0))\, d\tau(x_1 - x_0) \end{aligned}$$

y

$$\begin{aligned} \|F(x_1)\| &\leq \int_0^1 \|F'(x_0 + \tau(x_1 - x_0))\|\, d\tau\|x_1 - x_0\| \\ &\leq C\int_0^1 \tau\|x_1 - x_0\|\, d\tau\|x_1 - x_0\| \\ &\leq C\int_0^1 \tau(t_1 - t_0)(t_1 - t_0)\, d\tau \\ &= \int_0^1 (p'(t_0 + \tau(t_1 - t_0)) - p'(t_0))\,(t_1 - t_0)\, d\tau \\ &= \int_{t_0}^{t_1} (p'(s) - p'(t_0))\, ds \\ &= p(t_1) - p(t_0) - p'(t_0)(t_1 - t_0) \\ &= p(t_1), \end{aligned}$$

entonces

$$\|x_2 - x_1\| = \|[F'(x_1)]^{-1}F(x_1)\| \leq \|[F'(x_1)]^{-1}\|\|F(x_1)\| \leq -\frac{p(t_1)}{p'(t_1)} = t_2 - t_1,$$

$$\|x_2 - x_0\| \leq \|x_2 - x_1\| + \|x_1 - x_0\| \leq t_2 - t_1 + t_1 - t_0 = t_2 - t_0 < t^*.$$

y, en consecuencia, $x_2 \in B(x_0, t^*)$.

Procediendo ahora por inducción sobre n, se prueba fácilmente que la sucesión $\{x_n\}$ está bien definida, $x_n \in B(x_0, t^*)$, para todo $n \geq 0$, y

$$\|x_{n+1} - x_n\| = \|[F'(x_n)]^{-1}F(x_n)\| \leq \|[F'(x_n)]^{-1}\|\|F(x_n)\| \leq -\frac{p(t_n)}{p'(t_n)} = t_{n+1} - t_n,$$

$$\|x_{n+1} - x_0\| \leq \|x_{n+1} - x_n\| + \|x_n - x_0\| \leq t_{n+1} - t_n + t_n - t_0 = t_{n+1} - t_0 < t^*.$$

Como $\lim\limits_{n\to\infty} t_n = t^*$, entonces existe x^* tal que $x^* = \lim\limits_{n\to\infty} x_n$. Veamos que x^* es solución de $F(x) = 0$. De $\|F'(x_n) - F'(x_0)\| \leq C\|x_n - x_0\| \leq Ct^*$, se sigue $\|F'(x_n)\| \leq F'(x_0) + Ct^*$, para todo $n \geq 0$. Entonces, como $F(x_n) = F'(x_n)(x_{n+1} - x_n)$, pasando al límite cuando $n \to \infty$ y teniendo en cuenta que la sucesión $\{\|F'(x_n)\|\}$ está acotada, obtenemos $F(x^*) = 0$ por la continuidad de F.

Finalmente, probamos la unicidad de solución. Suponemos que existe otra solución $y^* \in B(x_0, t^{**}) \cap \Omega$ de la ecuación $F(x) = 0$ distinta de x^*. Consideramos

$$F(y^*) - F(x^*) = \int_{x^*}^{y^*} F'(x)\,dx = \int_0^1 F'(x^* + t(y^* - x^*))(y^* - x^*)\,dt = 0 \tag{1.39}$$

y el operador $J = \displaystyle\int_0^1 F'(x^* + t(z^* - x^*))\,dt$. Como

$$\begin{aligned}
\|I - [F'(x_0)]^{-1}J\| &\leq \|[F'(x_0)]^{-1}\|\|F'(x_0) - J\| \\
&\leq \|[F'(x_0)]^{-1}\|\|\int_0^1 \|F'(x^* + t(y^* - x^*)) - F'(x_0)\|\,dt \\
&\leq C\theta\int_0^1 \left((1-t)\|x^* - x_0\| + t(\|y^* - x_0\|)\right)\,dt \\
&< \frac{C\theta}{2}(t^* + t^{**}) \\
&= 1,
\end{aligned}$$

el operador J es inversible, por el lema de Banach (lema 1.22), y, por tanto, $x^* = y^*$. ■

Técnicas alternativas al principio de la mayorante para demostrar la convergencia semilocal de los métodos iterativos en espacios de Banach están basadas en relaciones de recurrencia. Desde que Kantorovich generalizó, a mediados del siglo XX, el método de Newton a espacios de Banach, varios autores han demostrado la convergencia semilocal del método de Newton mediante este tipo de técnicas, incluido el propio Kantorovich, cuyas primeras demostraciones están basadas en relaciones de recurrencia. Nosotros presentamos, a continuación, una técnica de este tipo, desarrollada por los autores del texto [20], que da muy buenos resultados y que es muy sencilla de utilizar.

En primer lugar, definimos la sucesión de números reales

$$a_0 = C\delta\theta^2, \qquad a_n = \frac{a_{n-1}^2}{2} f(a_{n-1})^2, \quad n \in \mathbb{N}, \quad \text{donde} \quad f(t) = \frac{1}{1-t}.$$

Notemos que consideraremos $a_0 > 0$, puesto que si $a_0 = 0$, queda un problema trivial por ser x_0 la solución de una ecuación $F(x) = 0$.

A continuación, mediante inducción matemática sobre n, se prueban, para $n \geq 1$, las siguientes dos relaciones de recurrencia entre la sucesión escalar $\{a_n\}$ y la dada por el método de Newton, $\{x_n\}$, en el espacio de Banach X:

(i_n) El operador $[F'(x_n)]^{-1}$ existe y es tal que $\|[F'(x_n)]^{-1}\| \leq f(a_{n-1})\|[F'(x_{n-1})]^{-1}\|$,

(ii_n) $\|x_{n+1} - x_n\| \leq \dfrac{a_{n-1}}{2} f(a_{n-1})\|x_n - x_{n-1}\|$.

Veamos que las dos relaciones de recurrencia anteriores se cumplen para $n = 1$. Suponiendo que $a_0 < 1$ y $x_1 \in \Omega$, se sigue

$$\|I - [F'(x_0)]^{-1}F'(x_1)\| \leq \|[F'(x_0)]^{-1}\|\|F'(x_0) - F'(x_1)\| \leq C\theta\|x_1 - x_0\| \leq C\delta\theta^2 = a_0 < 1.$$

Entonces, por el lema de Banach (lema 1.22), el operador $[F'(x_1)]^{-1}$ existe y es tal que

$$\|[F'(x_1)]^{-1}\| \leq \frac{\|[F'(x_0)]^{-1}\|}{1 - \|I - [F'(x_0)]^{-1}F'(x_1)\|} \leq f(a_0)\|[F'(x_0)]^{-1}\|.$$

Por tanto, tenemos (i_1).

Ahora, utilizando la fórmula de Taylor (teorema 1.38)

$$F(x_1) = \int_0^1 (F'(x_0 + \tau(x_1 - x_0)) - F'(x_0))\, d\tau(x_1 - x_0)$$

y (A3), obtenemos

$$\|F(x_1)\| \leq \frac{C}{2}\|x_1 - x_0\|^2 \leq \frac{C\delta\theta}{2}\|x_1 - x_0\|.$$

Además,

$$\|x_2 - x_1\| = \|[F'(x_1)]^{-1}F(x_1)\| \leq \|[F'(x_1)]^{-1}\|\|F(x_1)\| \leq \frac{a_0}{2} f(a_0)\|x_1 - x_0\|.$$

Por tanto, tenemos (ii_1).

Si suponemos ahora que $a_n < 1$ y $x_n \in \Omega$, para todo $n \in \mathbb{N}$, y que las relaciones (i_k)–(ii_k) se verifican para $k = 1, 2, \ldots, n$, las relaciones (i_{n+1})–(ii_{n+1}) se demuestran de forma totalmente análoga a las relaciones (i_1)–(ii_1) y se completa la inducción.

En segundo lugar, para probar la convergencia de la sucesión $\{x_n\}$ definida por el método de Newton en el espacio de Banach X, debemos probar que $\{x_n\}$ es una sucesión de Cauchy contenida en Ω y que $a_n < 1$, para todo $n \geq 0$. Para ello, comenzamos dando algunas propiedades de la sucesión escalar $\{a_n\}$ en el siguiente lema.

Lema 1.54. *Si* $a_0 < \frac{1}{2}$, *entonces*

(i) $\Delta = \dfrac{a_0}{2} f(a_0)^2 < 1$,

(ii) la sucesión $\{a_n\}$ *es estrictamente decreciente,*

(iii) $a_n < 1$, *para todo* $n \geq 0$.

Si $a_0 = \frac{1}{2}$, *entonces* $a_n = a_0 < 1$, *para todo* $n \in \mathbb{N}$.

Demostración. En primer lugar, consideramos el caso $a_0 < \frac{1}{2}$. El apartado (i) es inmediato. En cuanto a (ii), lo demostramos por inducción matemática sobre n. Como $\Delta < 1$, entonces $a_1 < a_0$. Suponemos ahora que $a_k < a_{k-1}$ para $k = 1, 2, \dots, n$. Luego,

$$a_{n+1} = \frac{a_n^2}{2} f(a_n)^2 < \Delta\, a_n < a_n,$$

por ser f creciente. Por tanto, la sucesión $\{a_n\}$ es estrictamente decreciente.

Para ver (iii), tenemos que $a_n < a_0 < 1$, para todo $n \geq 0$, por ser $\{a_n\}$ una sucesión estrictamente decreciente y $a_0 < \frac{1}{2}$.

Si, por otra parte, $a_0 = \frac{1}{2}$, entonces $\Delta = \dfrac{a_0}{2} f(a_0)^2 = 1$ y, en consecuencia, tenemos $a_n = a_0 = \frac{1}{2} < 1$, para todo $n \geq 0$. ■

A continuación, probamos el teorema de convergencia semilocal.

Teorema 1.55. *Sea $F : \Omega \subset X \to Y$ un operador una vez continuamente diferenciable Fréchet, definido en un dominio abierto convexo no vacío Ω de un espacio de Banach X y con valores en un espacio de Banach Y. Supongamos que se verifican las condiciones (A1)–(A3), $a_0 = C\delta\theta^2 \leq \frac{1}{2}$ y $B(x_0, \rho) \subset \Omega$, donde $\rho = \frac{2(1-a_0)}{2-3a_0}\delta\theta$. Entonces, la ecuación $F(x) = 0$ tiene una solución x^* y el método de Newton converge a x^* empezando en x_0. Además, $x_n, x^* \in \overline{B(x_0, \rho)}$, para todo $n \in \mathbb{N}$, y la solución x^* es única en la región $\overline{B\left(x_0, \frac{1}{C\theta}\right)} \cap \Omega$.*

Demostración. Sabemos por el lema anterior que $a_n < 1$ para todo $n \geq 0$. Veamos que $\{x_n\}$ es una sucesión de Cauchy y que $x_n \in B(x_0, \rho)$, para todo $n \geq 1$. Así, para $m \geq 1$ y $n \geq 1$, vemos

$$\begin{aligned} \|x_{n+m} - x_n\| &\leq \|x_{n+m} - x_{n+m-1}\| + \|x_{n+m-1} - x_{n+m-2}\| + \cdots + \|x_{n+1} - x_n\| \\ &\leq \left(1 + \sum_{i=n}^{n+m-2} \left(\prod_{j=n}^{i} \frac{a_j}{2} f(a_j) \right) \right) \|x_{n+1} - x_n\|. \end{aligned} \tag{1.40}$$

Como $\{a_j\}$ es estrictamente decreciente y f es creciente, por el lema anterior y (1.40), para $n \geq 1$, tenemos

$$\begin{aligned} \|x_{n+m} - x_n\| &< \sum_{i=n-1}^{n+m-2} \left(\prod_{j=0}^{i} \frac{a_0}{2} f(a_0) \right) \|x_1 - x_0\| \\ &= \sum_{i=n-1}^{n+m-2} \left(\frac{a_0}{2} f(a_0) \right)^{i+1} \|x_1 - x_0\| \\ &= \sum_{i=0}^{m-1} \left(\frac{a_0}{2} f(a_0) \right)^{i+n} \|x_1 - x_0\| \\ &= \Delta^n \frac{1 - \Delta^m}{1 - \Delta} \|x_1 - x_0\|. \end{aligned}$$

Luego, la sucesión $\{x_n\}$ es de Cauchy por ser $\Delta = \dfrac{a_0}{2} f(a_0)^2 < 1$.

Si ahora hacemos $n = 0$ en (1.40), entonces

$$\|x_m - x_0\| < \frac{1 - \Delta^m}{1 - \Delta} \|x_1 - x_0\| < \frac{\delta\theta}{1 - \Delta} = \rho.$$

Por lo tanto, $x_n \in B(x_0, \rho)$, para todo $n \in \mathbb{N}$, la sucesión $\{x_n\}$ está bien definida y es una sucesión de Cauchy. Luego, $\lim\limits_{n\to\infty} x_n = x^* \in \overline{B(x_0, \rho)}$. Veamos que x^* es una solución de $F(x) = 0$. Como $\|[F'(x_n)]^{-1}F(x_n)\| = \|x_{n+1} - x_n\| \to 0$ cuando $n \to \infty$, teniendo en cuenta

$$\|F(x_n)\| \leq \|F'(x_n)\|\|[F'(x_n)]^{-1}F(x_n)\|$$

y que la sucesión $\{\|F'(x_n)\|\}$ está acotada, puesto que

$$\|F'(x_n)\| \overset{\text{(A3)}}{\leq} \|F'(x_0)\| + C\|x_n - x_0\| < \|F'(x_0)\| + C\rho,$$

se sigue $\|F(x_n)\| \to 0$ cuando $n \to \infty$. En consecuencia, obtenemos $F(x^*) = 0$ por la continuidad de F en $\overline{B(x_0, \rho)}$.

Para probar la unicidad, suponemos que y^* es otra solución de $F(x) = 0$, distinta de x^*, en la región $\overline{B\left(x_0, \frac{1}{C\theta}\right)} \cap \Omega$. Entonces, a partir de la aproximación (1.39), tenemos que probar que el operador $J = \displaystyle\int_0^1 F'(x^* + t(y^* - x^*))\,dt$ es inversible. En efecto, como

$$\begin{aligned}
\|I - [F'(x_0)]^{-1}J\| &\leq \|[F'(x_0)]^{-1}\|\|F'(x_0) - J\| \\
&\leq \|[F'(x_0)]^{-1}\|\|\int_0^1 \|F'(x^* + t(y^* - x^*)) - F'(x_0)\|\,dt \\
&\leq C\theta \int_0^1 ((1-t)\|x^* - x_0\| + t(\|y^* - x_0\|))\,dt \\
&< \frac{C\theta}{2}\left(\rho + \tfrac{1}{C\theta}\right) \\
&= 1,
\end{aligned}$$

por el lema de Banach (lema 1.22), el operador J es inversible. ∎

Terminamos esta sección recordando que, utilizando la definición 1.52 de R-orden de convergencia, es conocido que el método de Newton tiene R-orden de convergencia al menos dos [27].

1.5.2. Accesibilidad del método de Newton

Tal y como hemos indicado anteriormente, podemos observar la accesibilidad de un método iterativo desde tres puntos de vista, dos de los cuales, la cuenca de atracción y la región de accesibilidad, son experimentales, y el otro, el dominio de parámetros, es teórico. Tanto la cuenca de atracción como la región de accesibilidad están asociadas a la ecuación a resolver, mientras que el dominio de parámetros no, ya que éste está asociado a un resultado de convergencia semilocal.

Para un sencillo entendimiento de estas tres formas de observar la accesibilidad de un método iterativo cuando se aplica a la resolución de una ecuación, vamos a utilizar el método de Newton y la ecuación compleja académica $F(z) = z^3 - 1 = 0$, donde $F : \mathbb{C} \to \mathbb{C}$. Notemos que la ecuación anterior tiene tres raíces complejas: $z^* = 1$, $z^{**} = \exp\left(\frac{2\pi i}{3}\right)$ y $z^{***} = \exp\left(-\frac{2\pi i}{3}\right)$.

Cuenca de atracción

Recordamos que la cuenca de atracción es el conjunto de todos los puntos de salida a partir de los cuales un método iterativo converge a una solución una vez fijada una tolerancia o un número máximo de iteraciones. En la figura 1.3 mostramos las cuencas de atracción asociadas a las tres raíces de la ecuación anterior cuando aplicamos el método de Newton para su aproximación. Para representar las cuencas de atracción hemos considerado un rectángulo D en el plano complejo $\mathbb{C}$ y asignado un color para cada raíz a la cual el método de Newton converge.

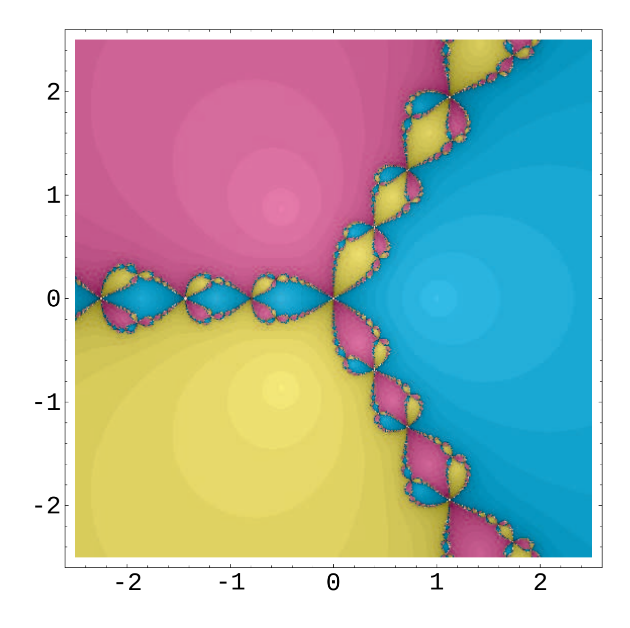


Figura 1.3: Cuencas de atracción de las tres raíces de $z^3 - 1 = 0$ cuando se aproximan mediante el método de Newton.

En la práctica, consideramos una malla de 512×512 puntos en D y elegimos estos puntos como x_0. Utilizamos el rectángulo $[-2.5, 2.5] \times [-2.5, 2.5]$ que contiene a las tres raíces. El método de Newton, empezando en un punto $x_0 \in D$, puede converger a cualquiera de las tres raíces o, eventualmente, diverger. En todos los casos, utilizamos la tolerancia 10^{-3} y un máximo de 50 iteraciones. Si no obtenemos la tolerancia deseada con 50 iteraciones, no se continúa y decidimos que el método iterativo no converge a ninguna raíz comenzando en z_0. En particular, hemos utilizado respectivamente los colores cyan, magenta y amarillo para las cuencas de atracción de las raíces z^*, z^{**} y z^{***}, respectivamente. El color se hace más claro o más oscuro según sea el número de iteraciones utilizadas para alcanzar una raíz con la precisión fijada. No hemos pintado los puntos del rectángulo correspondientes a las aproximaciones iniciales a partir de las cuales no se alcanza ninguna de las raíces con tolerancia 10^{-3} en un máximo de 50 iteraciones. Para otras estrategias se puede consultar [47] y las referencias allí dadas. Los gráficos se han generado con Mathematica 5.1 [49].

Región de accesibilidad

Sabemos que los puntos de salida del método de Newton tienen asociados los parámetros δ, θ y C dados en las condiciones iniciales (A1)–(A3). Para representar las regiones de acce-

sibilidad del método de Newton, coloreamos los puntos cuyos parámetros asociados verifican las condiciones de convergencia y, en otro caso, no los coloreamos. La región de accesibilidad asociada a una solución de una ecuación nos indica entonces el dominio de puntos de salida a partir de los cuales tenemos asegurada la convergencia del método de Newton (es decir, el conjunto de puntos de salida que satisfacen las condiciones de convergencia para el método de Newton).

A continuación, en la figura 1.4, vemos cuáles son las regiones de accesibilidad asociadas a las raíces de la ecuación $z^3 - 1 = 0$ cuando se aproximan mediante el método de Newton. Representamos la regiones de accesibilidad coloreando los puntos x_0 que verifican la condición $C\delta\theta^2 \leq \frac{1}{2}$ del teorema 1.53.

Hemos utilizado los colores cyan, magenta y amarillo para regiones de accesibilidad de las raíces z^*, z^{**} y z^{***}, respectivamente. Y, al igual que para las cuencas de atracción, el color se hace más claro o más oscuro según sea el número de iteraciones utilizadas para alcanzar una solución con la precisión fijada. Los gráficos se han generado de nuevo con Mathematica 5.1 [49].

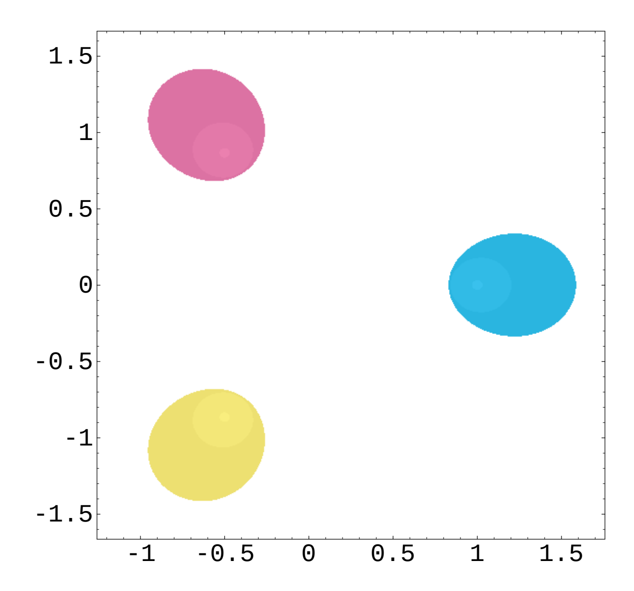


Figura 1.4: Regiones de accesibilidad de las tres raíces de la ecuación $z^3 - 1 = 0$ para el método de Newton según el teorema de Newton-Kantorovich (teorema 1.53).

Dominio de parámetros

Aparte de la observación empírica de la accesibilidad del método de Newton dado por las cuencas de atracción y las regiones de accesibilidad, podemos realizar un estudio de la accesibilidad de dicho método a partir de las condiciones de convergencia impuestas en el teorema de convergencia semilocal 1.53.

Observamos que las condiciones de convergencia impuestas en el teorema 1.53 tienen dos partes diferenciadas. Por una parte, las condiciones iniciales (A1)–(A2) exigidas al punto de salida x_0; y por otra, la condición (A3) exigida al operador F. Para estudiar las restricciones que se imponen con las condiciones iniciales, utilizamos el dominio de parámetros, que esta-

blece gráficamente en un plano real la relación entre los parámetros que se definen a partir de las condiciones iniciales.

A partir del teorema 1.53, si queremos estudiar teóricamente la accesibilidad del método de Newton, basta con tener en cuenta que, dado $x_0 \in \Omega$, el método tiene asociados los parámetros δ y θ que aparecen en (A1) y (A2). Así, a partir de la condición de convergencia $C\delta\theta^2 \leq \frac{1}{2}$ del teorema 1.53, podemos definir el dominio de parámetros asociado a dicho teorema como el conjunto del plano real dado por $\{(\delta,\theta) \in \mathbb{R}^2 : \ C\delta\theta^2 \leq \frac{1}{2}\}$. Por un lado, observamos que el parámetro δ mide la aproximación de x_0 a la solución x^*. Notemos entonces que $\delta = 0$ si $x_0 = x^*$. Por otro lado, observamos que el parámetro C, que se define a partir de la condición (A3) que se exige al operador F, es siempre una cantidad fija, de manera que no va a influir en el dominio de parámetros.

En la figura 1.5 mostramos el dominio de parámetros del método de Newton asociado al teorema 1.53. Para representarlo gráficamente, hemos considerado el plano xy con $x = \theta$ (eje de abscisas) e $y = C\delta$ (eje de ordenadas) y coloreado en rojo los valores de los parámetros que verifican la condición $x^2y \leq \frac{1}{2}$.

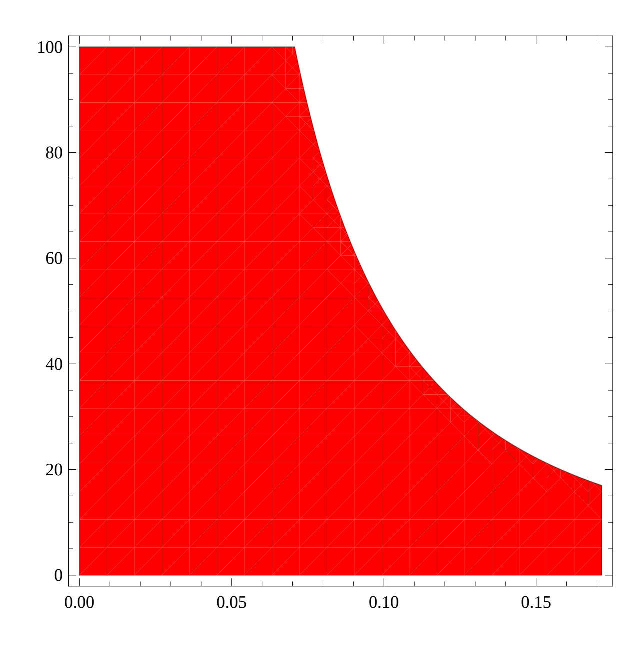


Figura 1.5: Dominio de parámetros del método de Newton asociado al teorema de Newton-Kantorovich (teorema 1.53).

1.6. Algunas ecuaciones no lineales en espacios de Banach

La resolución de sistemas de ecuaciones no lineales de la forma $F(x) = 0$, donde $F : D \subset \mathbb{R}^m \to \mathbb{R}^m$ es una función definida en un dominio abierto convexo no vacío D, es un problema habitual de las ciencias y la ingeniería. Es importante entonces destacar que los conocidos esquemas en diferencias finitas permiten transformar problemas continuos, como ecuaciones integrales y ecuaciones diferenciales, en sistemas de ecuaciones, tal como vemos

a continuación, puesto que uno de los objetivos de este texto se centra en la resolución de sistemas de ecuaciones no lineales en $\mathbb{R}^m$.

1.6.1. Ecuaciones integrales de Hammerstein

Las ecuaciones de Hammerstein tienen un origen físico importante y surgen de la dinámica de fluidos electromagnéticos [39]. En particular, las ecuaciones integrales no lineales de tipo Hammerstein mixto son de la forma:

$$x(s) = f(s) + \int_a^b G(s,t)H(t,x(t))\,dt, \quad s \in [a,b], \tag{1.41}$$

donde $-\infty < a < b < +\infty$, $f(s)$ es una función continua dada en $[a,b]$ y el núcleo G y la función H son conocidos. Estas ecuaciones aparecieron a principios de los años 30 del siglo XX como modelos generales del estudio de problemas de valores en la frontera semilineales, donde el núcleo $G(s,t)$ se presenta típicamente como la función de Green de un operador diferencial [21]. Así, la ecuación (1.41) se puede reformular como un problema de valores en la frontera de dos puntos con una cierta condición de contorno no lineal [6]. También aparecen análogos multidimensionales de la ecuación (1.41) como reformulaciones de una EDP elíptica con condiciones de contorno no lineales [33]. Ecuaciones integrales como (1.41) aparecen frecuentemente en numerosas aplicaciones del mundo real [9]. Por ejemplo, algunos problemas considerados en la teoría vehicular, la biología y la teoría de colas llevan a ecuaciones integrales de este tipo [17]. Estas ecuaciones también se aplican en la teoría de la transferencia radiactiva, en la teoría del transporte de neutrones y en la teoría cinética de gases [25]. Destacamos además el papel significativo que juegan en varias aplicaciones [14], como por ejemplo los modelos dinámicos de reactores químicos [12], que están gobernados por ecuaciones de control, justificando así su estudio y resolución [22].

La resolución de la ecuación integral (1.41) es equivalente a resolver la ecuación $\mathcal{F}(x) = 0$, donde $\mathcal{F} : \Omega \subset \mathcal{C}([a,b]) \to \mathcal{C}([a,b])$ y

$$[\mathcal{F}(x)](s) = x(s) - f(s) - \int_a^b G(s,t)H(t,x(t))\,dt, \quad s \in [a,b]. \tag{1.42}$$

Notemos que, como ya hemos visto, $\mathcal{C}([a,b])$ es un espacio de Banach con la norma $\|\cdot\|_\infty$ y, por tanto, el operador (1.42) está definido entre dos espacios de Banach.

Cuando queremos aproximar una solución de la ecuación $F(\mathbf{x}) = 0$, donde el operador F está definido por $F : D \subseteq \mathbb{R}^m \to \mathbb{R}^m$ en un dominio abierto convexo no vacío D, mediante el método de Newton

$$\begin{cases} \mathbf{x_0} \in D, \\ \mathbf{x_{n+1}} = \mathbf{x_n} - [F'(\mathbf{x_n})]^{-1}F(\mathbf{x_n}), \quad n \geq 0, \end{cases}$$

lo que hacemos es resolver el sistema de ecuaciones lineales en cada paso dado por

$$F'(\mathbf{x_n})(\mathbf{x_{n+1}} - \mathbf{x_n}) = -F(\mathbf{x_n}). \tag{1.43}$$

En cambio, si el operador es de la forma $\mathcal{F} : \mathcal{C}([a,b]) \to \mathcal{C}([a,b])$, no podemos utilizar la idea anterior porque no sabemos resolver la ecuación integral que corresponde a la ecuación (1.43) a partir de (1.42). Tampoco podemos aplicar directamente el método de Newton ya que no conocemos el operador $[\mathcal{F}'(x)]^{-1}$.

Así, como primer paso, discretizamos la ecuación (1.41) para transformarla en un problema de dimensión finita. Consideramos entonces (1.41) siendo el núcleo G la función de Green en $[a,b]\times[a,b]$ y aproximamos la integral que aparece en (1.41) usando la siguiente fórmula de cuadratura numérica de Gauss-Legendre con m nodos [28]

$$\int_a^b q(t)\,dt \simeq \sum_{i=1}^m w_i q(t_i),$$

donde los nodos t_i y los pesos w_i son conocidos.

Si denotamos la aproximación de $x(t_i)$ por x_i y la de $f(t_i)$ por f_i $(i=1,2,\ldots,m)$, entonces la ecuación (1.41) se transforma en el siguiente sistema de ecuaciones no lineales

$$x_i = f_i + \sum_{j=1}^m a_{ij} H(t_j, x_j)\,, \quad j = 1,2,\ldots,m, \tag{1.44}$$

donde

$$a_{ij} = w_j\, G(t_i,t_j) = \begin{cases} w_j \, \frac{(b-t_i)(t_j-a)}{b-a} & \text{si } j \le i, \\ w_j \, \frac{(b-t_j)(t_i-a)}{b-a} & \text{si } j > i. \end{cases}$$

Ahora, el sistema (1.44) se puede escribir como

$$F(\mathbf{x}) \equiv \mathbf{x} - \mathbf{f} - A\,\mathbf{y} = 0, \qquad F : \mathbb{R}^m \longrightarrow \mathbb{R}^m, \tag{1.45}$$

donde

$$\mathbf{x} = (x_1,x_2,\ldots,x_m)^T, \qquad \mathbf{f} = (f_1,f_2,\ldots,f_m)^T, \qquad A = (a_{ij})_{i,j=1}^m,$$
$$\mathbf{y} = (H(t_1,x_1), H(t_2,x_2), \ldots, H(t_m,x_m))^T.$$

Por otra parte, como los métodos iterativos que estudiamos en este texto utilizan diferencias divididas de primer orden en su algoritmo y en $\mathbb{R}^m$ podemos considerar diferencias divididas de primer orden que no necesitan que la función sea diferenciable [37], consideraremos la diferencia dividida de primer orden dada por $[\mathbf{u},\mathbf{v};F] = ([\mathbf{u},\mathbf{v};F]_{ij})_{i,j=1}^m \in \mathcal{L}(\mathbb{R}^m,\mathbb{R}^m)$, con

$$[\mathbf{u},\mathbf{v};F]_{ij} = \frac{1}{u_j - v_j}\left(F_i(u_1,\ldots,u_j,v_{j+1},\ldots,v_m) - F_i(u_1,\ldots,u_{j-1},v_j,\ldots,v_m)\right), \tag{1.46}$$

$\mathbf{u} = (u_1,u_2,\ldots,u_m)^T$ y $\mathbf{v} = (v_1,v_2,\ldots,v_m)^T$. Así, para la función F definida en (1.45), tenemos $[\mathbf{u},\mathbf{v};F] = I - A\,\mathrm{diag}\{\mathbf{z}\}$, donde $\mathbf{z} = (z_1,z_2,\ldots,z_m)^T$ y $z_i = \dfrac{H(t_i,u_i) - H(t_i,v_i)}{u_i - v_i}$, para todo $i = 1,2,\ldots,m$.

1.6.2. Problemas conservativos

Es bien conocido que la energía se disipa en la acción de cualquier sistema dinámico real, generalmente a través de algún tipo de fricción. Sin embargo, en ciertas situaciones esta disipación es tan lenta que se puede despreciar en periodos de tiempo relativamente cortos. En tales casos se supone la ley de conservación de la energía, es decir, que la suma de la energía cinética y la energía potencial sea constante. Un sistema de este tipo se dice que es conservativo.

Si φ y ψ son funciones arbitrarias con la propiedad de que $\varphi(0) = 0$ y $\psi(0) = 0$, la ecuación general

$$\mu \frac{d^2x(t)}{dt^2} + \psi\left(\frac{dx(t)}{dt}\right) + \varphi(x(t)) = 0 \tag{1.47}$$

se puede interpretar como la ecuación del movimiento de una masa μ bajo la acción de una fuerza restauradora $-\psi\left(\frac{dx}{dt}\right)$. En general, estas fuerzas no son lineales y la ecuación (1.47) se puede considerar como una ecuación básica de mecánica no lineal. Ahora vamos a considerar el caso especial de un sistema no lineal conservativo descrito por la ecuación

$$\mu \frac{d^2x(t)}{d^2t} + \varphi(x(t)) = 0, \tag{1.48}$$

en la que la fuerza de amortiguación es nula y, en consecuencia, no hay disipación de energía. Diversos estudios de (1.47), con aplicaciones a un gran número de problemas físicos, se pueden encontrar en las referencias clásicas [3] y [44].

Ahora, consideramos el caso especial de un sistema no lineal conservativo descrito por la ecuación

$$\frac{d^2x(t)}{dt^2} + \phi(x(t)) = 0 \tag{1.49}$$

con condiciones de contorno

$$x(0) = x(1) = 0. \tag{1.50}$$

La resolución de la ecuación diferencial (1.49) es equivalente a resolver la ecuación $\mathcal{F}(x) = 0$, donde $\mathcal{F} : \mathcal{C}^2([0,1]) \to \mathcal{C}([0,1])$ y

$$[\mathcal{F}(x)](t) = \frac{d^2x(t)}{dt^2} + \phi(x(t)).$$

Notemos que, como ya hemos visto, $\mathcal{C}^2([0,1])$ es un espacio de Banach con la norma $\|x\|_\infty = \text{máx}\{\|x\|_\infty, \|x'\|_\infty\}$ teniendo en cuenta que $\mathcal{C}([0,1])$ es un espacio de Banach con la norma $\|\cdot\|_\infty$, de manera que el operador anterior $\mathcal{F}$ está definido entre dos espacios de Banach.

Tal y como hemos indicado anteriormente, estamos interesados en aproximar una solución de una ecuación no lineal $F(x) = 0$, donde F es un operador definido en un dominio abierto convexo no vacío D de $\mathbb{R}^m$ y tal que $F : D \subset \mathbb{R}^m \to \mathbb{R}^m$. Así, a continuación, utilizamos un proceso de discretización para transformar el problema de contorno de segundo orden en un problema finito-dimensional. Transformamos así el problema (1.49)–(1.50) en un sistema de ecuaciones no lineales. Para ello, aproximamos la segunda derivada por una fórmula numérica estándar.

En primer lugar, introducimos los puntos $t_j = jh$, $j = 0, 1, \ldots, m+1$, donde $h = \frac{1}{m+1}$ y m es un entero apropiado. El esquema es entonces designado por la determinación de los números x_j y se espera aproximar los valores $x(t_j)$ de la solución exacta en los puntos t_j. Una aproximación estándar para la segunda derivada en estos puntos es

$$x_j'' \approx \frac{x_{j-1} - 2x_j + x_{j+1}}{h^2}, \quad j = 1, 2, \ldots, m,$$

de manera que un procedimiento natural para obtener dicho esquema es exigir que los x_j satisfagan en cada punto t_j del interior de la malla la ecuación diferencial y, por la aproximación indicada, tenemos

$$x_{j-1} - 2x_j + x_{j+1} + h^2\phi(x_j) = 0, \tag{1.51}$$

y como x_0 y x_{m+1} están determinados por las condiciones de contorno, las incógnitas son $x_1, x_2, \ldots, x_m$.

Adicionalmente, simplificamos mediante el uso de notación vectorial y matricial. Introducimos entonces los vectores

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \qquad v_{\mathbf{x}} = \begin{pmatrix} \phi(x_1) \\ \phi(x_2) \\ \vdots \\ \phi(x_m) \end{pmatrix}$$

y la matriz

$$A = \begin{pmatrix} -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \cdots & 0 \\ 0 & 1 & -2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -2 \end{pmatrix},$$

de manera que el sistema de ecuaciones, que surge de imponer que (1.51) se verifique para $j = 1, 2, \ldots, m$, se puede escribir como

$$F(\mathbf{x}) \equiv A\mathbf{x} + h^2 v_{\mathbf{x}} = 0, \tag{1.52}$$

donde F es una función de $\mathbb{R}^m$ en $\mathbb{R}^m$.

Por otra parte, tal y como hemos dicho antes, al estudiar en este texto métodos iterativos que utilizan diferencias divididas de primer orden en su algoritmo, consideraremos la diferencia dividida de primer orden dada por $[\mathbf{u}, \mathbf{v}; F] = ([\mathbf{u}, \mathbf{v}; F]_{ij})_{i,j=1}^m \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$, con $[\mathbf{u}, \mathbf{v}; F]_{ij}$ definida en (1.46), de manera que en este caso

$$[\mathbf{u}, \mathbf{v}; F] = A + h^2 \mathrm{diag}\{\mathbf{z}\},$$

donde $\mathbf{z} = (z_1, z_2, \ldots, z_m)^T$ y $z_i = \dfrac{\phi(u_i) - \phi(v_i)}{u_i - v_i}$, para todo $i = 1, 2, \ldots, m$.

Terminamos diciendo que, a lo largo de todo el texto, por una parte, denotamos

$$\overline{B(x, \varrho)} = \{y \in X; \|y - x\| \leq \varrho\} \qquad \text{y} \qquad B(x, \varrho) = \{y \in X; \|y - x\| < \varrho\},$$

y, por otra parte, suponemos que existen todas las diferencias divididas de primer orden para cada par de puntos distintos del espacio de Banach X.

Parte II

MÉTODOS TIPO SECANTE

Como ya se ha indicado en la introducción de este texto, uno de nuestros objetivos principales es el estudio de métodos iterativos que no utilizan derivadas en su algoritmo. En general, estos métodos iterativos tienen el inconveniente de que no es sencillo localizar puntos de salida a partir de los cuales se asegure la convergencia de los mismos.

En esta segunda parte del texto, centramos nuestra atención en una familia de métodos iterativos con memoria, la de los tipo secante que, en el caso escalar, viene dada por el algoritmo

$$\begin{cases} \text{dados } t_{-1} \text{ y } t_0, \\ s_n = \lambda t_n + (1-\lambda) t_{n-1}, \quad \lambda \in [0,1], \\ t_{n+1} = t_n - \dfrac{s_n - t_n}{f(s_n) - f(t_n)} f(t_n), \quad n \geq 0, \end{cases}$$

cuando se aplica a la ecuación escalar $f(t) = 0$. Esta familia surge a partir de las interpretaciones geométricas del método de la secante y el método de Newton. Una característica importante de esta familia es que no utiliza derivadas en su algoritmo. Además, tiene R-orden de convergencia al menos superlineal, $\frac{1+\sqrt{5}}{2}$, y a medida que vamos considerando mayores valores de λ, próximos a uno, la velocidad de convergencia aumenta. Notemos que para $\lambda \to 1$ se obtiene el método de Newton, cuyo R-orden de convergencia es al menos cuadrático.

Tal y como hemos dicho anteriormente, son muchos los problemas de las ciencias y la ingeniería cuya resolución pasa por considerar el problema de aproximar una raíz x^* de una ecuación

$$F(x) = 0.$$

Para contemplar una mayor generalidad de la ecuación anterior, vamos a considerar que F es un operador definido en un subconjunto abierto convexo no vacío Ω de un espacio de Banach X y con valores en un espacio de Banach Y. En estas condiciones tan generales, la ecuación $F(x) = 0$ puede representar una ecuación escalar, un sistema de ecuaciones, una ecuación diferencial, una ecuación integral, etc.

Comenzamos extendiendo la familia de métodos iterativos tipo secante anterior a espacios de Banach con el objetivo de aproximar una solución x^* de la ecuación $F(x) = 0$, que queda de la siguiente forma:

$$\begin{cases} \text{dados } x_{-1}, x_0 \text{ en } \Omega, \\ y_n = \lambda x_n + (1-\lambda) x_{n-1}, \quad \lambda \in [0,1), \\ x_{n+1} = x_n - [y_n, x_n; F]^{-1} F(x_n), \quad n \geq 0, \end{cases}$$

donde $[x, y; F]$ es un operador diferencia dividida de primer orden de F en los puntos x e y.

Por una parte, tenemos que para $\lambda = 0$, obtenemos el método de la secante:

$$\begin{cases} \text{dados } x_{-1}, x_0 \text{ en } \Omega, \\ x_{n+1} = x_n - [x_{n-1}, x_n; F]^{-1} F(x_n), \quad n \geq 0, \end{cases}$$

cuyo R-orden de convergencia es al menos superlineal, $\frac{1+\sqrt{5}}{2}$. Por otra parte, si $\lambda = 1$ y el operador F es diferenciable, entonces $y_n = x_n$, $[y_n, x_n; F] = F'(x_n)$ y obtenemos el método de Newton:

$$\begin{cases} \text{dado } x_0 \text{ en } \Omega, \\ x_{n+1} = x_n - [F'(x_n)]^{-1} F(x_n), \quad n \geq 0. \end{cases}$$

Aunque el método de la secante es menos utilizado que el método de Newton, su utilización tiene gran interés puesto que no requiere la evaluación del operador derivada primera de F.

En esta segunda parte del texto mostramos el principal problema que tiene el resultado de convergencia semilocal dado para la familia de métodos iterativos tipo secante que se obtiene mediante una técnica basada en relaciones de recurrencia [24]. Es conocido que las hipótesis iniciales de todo resultado de convergencia semilocal para métodos iterativos tienen dos partes diferenciadas. Por una parte, las condiciones iniciales exigidas a los puntos de salida; y por otra, las condiciones exigidas al operador F. Pues bien, aquí nos ocuparemos de analizar las condiciones iniciales para mejorar un resultado de convergencia semilocal dado a partir de relaciones de recurrencia. Para estudiar las restricciones que imponen las condiciones iniciales utilizaremos la región de accesibilidad y el dominio de parámetros. El problema principal que presenta el resultado de convergencia semilocal dado a partir de relaciones de recurrencia corresponde con la situación que se plantea habitualmente cuando aplicamos los métodos tipo secante para aproximar una solución de $F(x) = 0$: no es sencillo localizar puntos de salida a partir de los cuales se asegure la convergencia de los métodos tipo secante.

Resolvemos el problema anterior mediante dos procedimientos. En el capítulo 2, para un operador diferenciable F, utilizamos un método iterativo híbrido de tipo predictor-corrector, que es un esquema iterativo que facilita la aplicación de un método iterativo, llamado corrector, a partir de la aplicación de otro método iterativo, llamado predictor, que permite localizar puntos de salida a partir de los cuales la convergencia del método corrector está asegurada. La idea básica es aplicar, hasta una cierta aproximación N_0, un método iterativo con R-orden de convergencia bajo, pero con buen dominio de puntos de salida, y utilizar después esta iteración como punto inicial para un método iterativo con mayor R-orden de convergencia. La clave está en el valor N_0, que juega un papel fundamental en la construcción de estos métodos iterativos híbridos. En concreto, en el capítulo 2, utilizamos un método iterativo híbrido (predictor-corrector) que facilita la aplicabilidad de los métodos tipo secante a partir del método simplificado de la secante (convergencia lineal), permitiendo así localizar puntos de salida, a partir de los cuales esté asegurada la convergencia de los métodos iterativos tipo secante, y aprovechar entonces su convergencia superlineal. En el capítulo 3, obraremos de forma diferente y utilizaremos una modificación de la técnica basada en relaciones de recurrencia para obtener un resultado de convergencia semilocal para la familia de métodos tipo secante, que es menos exigente a la hora de obtener puntos de salida adecuados para estos métodos. Además, el nuevo resultado de convergencia semilocal que se obtiene tiene la ventaja de que se puede aplicar a la resolución de ecuaciones en las que el operador implicado F es tanto diferenciable como no diferenciable.

Capítulo 2

Situación diferenciable

En este capítulo, como ya hemos indicado, teniendo en cuenta la deficiente accesibilidad que presentan los métodos iterativos que no utilizan derivadas y, en particular, los métodos tipo secante, nos planteamos mejorar su accesibilidad. Para conseguir este objetivo, planteamos la construcción de un método iterativo híbrido (predictor-corrector).

En la sección 2.1.1 presentamos un conocido resultado de convergencia semilocal para los métodos tipo secante, lo analizamos mediante regiones de accesibilidad y el dominio de parámetros asociado y vemos cuáles son las deficiencias que presenta. En la sección 2.2.1, una vez detectado correctamente el problema de la deficiente accesibilidad de los métodos tipo secante, a partir del resultado de convergencia semilocal presentado, consideramos, para resolverlo, el método simplificado de la secante en espacios de Banach,

$$\begin{cases} \text{dados } z_{-1}, z_0 \text{ en } \Omega, \\ z_{n+1} = z_n - [z_{-1}, z_0; F]^{-1} F(z_n), \quad n \geq 0, \end{cases} \tag{2.1}$$

cuyo R-orden de convergencia es al menos lineal, con el objetivo principal de construir un método iterativo híbrido (predictor-corrector) que utilice el método simplificado de la secante (2.1) como predictor y la familia de métodos tipo secante como corrector. Para ello, analizamos la convergencia semilocal del método simplificado de la secante (2.1) y, como las condiciones de convergencia impuestas a este método son menos restrictivas que las impuestas previamente a los métodos tipo secante, vemos que se pueden mejorar las regiones de accesibilidad y los dominios de parámetros de los métodos tipo secante a partir del método simplificado de la secante (2.1). Así, en la sección 2.3, mediante el método iterativo híbrido (predictor-corrector) construído, garantizamos la convergencia de los métodos tipo secante saliendo desde los mismos puntos de salida a partir de los cuales está garantizada la convergencia del método simplificado de la secante (2.1). Finalmente, en la sección 2.4, ilustramos todo lo anterior con la resolución de un sistema de ecuaciones no lineales.

2.1. Método corrector: los métodos tipo secante

2.1.1. Convergencia semilocal

En esta sección presentamos un resultado de convergencia semilocal para los métodos tipo secante y en el que la técnica de demostración está basada en relaciones de recurrencia. Para ello, suponemos que se cumplen las siguientes condiciones iniciales:

(C1) $\|x_0 - x_{-1}\| = \alpha \neq 0$ con $x_{-1}, x_0 \in \Omega$,

(C2) fijado $\lambda \in [0,1)$, existe $A_0^{-1} = [y_0, x_0; F]^{-1} \in \mathcal{L}(Y, X)$, para $x_0, y_0 \in \Omega$, y es tal que que $\|A_0^{-1}\| \leq \beta$,

(C3) $\|A_0^{-1} F(x_0)\| \leq \eta$,

(C4) $\|[x, y; F] - [u, v; F]\| \leq K(\|x - u\| + \|y - v\|)$; $K \geq 0$; $x, y, u, v \in \Omega$; $x \neq y$, $u \neq v$.

En primer lugar, se definen las siguientes sucesiones de números reales positivos:

$$a_n = f(a_{n-1}) g(a_{n-1}) b_{n-1}, \quad b_n = f(a_{n-1})^2 a_{n-1} b_{n-1}, \quad n \geq 0, \tag{2.2}$$

donde

$$a_{-1} = \frac{\eta}{\alpha + \eta}, \quad a_0 = f(a_{-1}) g(a_{-1}) b_{-1}, \quad b_{-1} = \frac{K \beta \alpha^2}{\alpha + \eta}, \tag{2.3}$$

$$f(t) = \frac{1}{1-t}, \quad g(t) = (1 - \lambda) + (1 + \lambda) f(t) t, \quad \lambda \in [0,1).$$

Notemos que tanto $f(t)$ como $g(t)$ son funciones crecientes en $\mathbb{R} - \{1\}$ y, además, $f(t) > 1$ en $(0,1)$.

A partir de las condiciones iniciales (C1)–(C4), si x_1 está bien definido, se deduce que existe $A_0^{-1} = [y_0, x_0; F]^{-1}$ y

$$\begin{aligned} \|x_1 - x_0\| = \|A_0^{-1} F(x_0)\| \leq \eta = f(a_{-1}) a_{-1} \|x_0 - x_{-1}\|, \\ K\|x_1 - x_0\| \, \|x_0 - x_{-1}\| \leq K \beta \alpha = f(a_{-1}) b_{-1}. \end{aligned} \tag{2.4}$$

A continuación, se prueban, mediante inducción matemática sobre n, las siguientes tres relaciones de recurrencia para $n \geq 1$:

(i_n) Existe $A_n^{-1} = [y_n - x_n; F]^{-1}$ y es tal que $\|A_n^{-1}\| \leq f(a_{n-1}) \|A_{n-1}^{-1}\|$,

(ii_n) $\|x_{n+1} - x_n\| \leq f(a_{n-1}) a_{n-1} \|x_n - x_{n-1}\|$,

(iii_n) $K\|A_n^{-1}\| \|x_n - x_{n-1}\| \leq f(a_{n-1}) b_{n-1}$.

Veámoslo para $n = 1$. Suponiendo que $a_0 < 1$ y $x_1 \in \Omega$, de (2.3) y (2.4), se sigue:

$$\begin{aligned} \|I - A_0^{-1} A_1\| &\leq \|A_0^{-1}\| \|A_0 - A_1\| \\ &\leq \|A_0^{-1}\| \left(\|y_1 - y_0\| + \|x_1 - x_0\| \right) \\ &\leq K \|A_0^{-1}\| \left[(1 - \lambda) + (1 + \lambda) f(a_{-1}) a_{-1} \right] \|x_0 - x_{-1}\| \\ &\leq a_0 \\ &< 1. \end{aligned}$$

Entonces, por el lema de Banach (lema 1.22), existe A_1^{-1} y es tal que

$$\|A_1^{-1}\| \leq f(a_0) \|A_0^{-1}\|.$$

Por tanto, se cumple (i_1).

Ahora, usando la fórmula de Taylor (teorema 1.38)

$$F(x_1) = (F'(x_0) - A_0)(x_1 - x_0) + \int_0^1 (F'(x_0 + t(x_1 - x_0)) - F'(x_0))(x_1 - x_0)dt,$$

y teniendo en cuenta que F es diferenciable por cumplirse (C4), lema 1.43, etonces $[x, x; F] = F'(x)$ y

$$\begin{aligned} \|F(x_1)\| &\leq K((1-\lambda)\|x_0 - x_{-1}\| + \|x_1 - x_0\|)\|x_1 - x_0\| \\ &\leq Kg(a_{-1})\|x_0 - x_{-1}\|\|x_1 - x_0\|. \end{aligned}$$

Como A_1^{-1} existe, si x_2 está bien definido, se sigue

$$\|x_2 - x_1\| \leq f(a_0)\|A_0^{-1}\|\|F(x_1)\| \leq f(a_0)a_0\|x_1 - x_0\|$$

y tenemos (ii_2).

Notemos que, como consecuencia de (2.4) y (i_1), se sigue (iii_1), puesto que

$$K\|A_1^{-1}\|\|x_1 - x_0\| \leq Kf(a_0)\|A_0^{-1}\|\|x_1 - x_0\| \leq f(a_0)b_0.$$

Finalmente, suponiendo $a_n < 1$ y $x_n \in \Omega$, para todo $n \geq 1$, el paso inductivo (i_{n+1})–(iii_{n+1}) se demuestra de forma totalmente análoga y se completa la inducción.

En segundo lugar, para probar la convergencia de la sucesión $\{x_n\}$ que definen los métodos tipo secante, se estudian las sucesiones reales definidas en (2.2). Debemos probar que $\{x_n\}$ es una sucesión de Cauchy contenida en Ω y que $a_n < 1$, para todo $n \geq 0$.

Se empieza denotando la sucesión de Fibonacci por $\{\delta_n\}$, que se define como sigue

$$\delta_1 = \delta_2 = 1 \quad y \quad \delta_{n+2} = \delta_{n+1} + \delta_n, \quad n \geq 1,$$

$$s_n = \delta_1 + \delta_2 + \cdots + \delta_n, \quad n \geq 1,$$

y se demuestran por inducción las siguientes dos propiedades:

- $\delta_n = \dfrac{1}{\sqrt{5}}\left[\left(\dfrac{1+\sqrt{5}}{2}\right)^n - \left(\dfrac{1-\sqrt{5}}{2}\right)^n\right] > \dfrac{1}{\sqrt{5}}\left(\dfrac{1+\sqrt{5}}{2}\right)^{n-1}, \quad n \geq 1.$

- $s_n = \delta_{n+2} - 1$ y $\mu_n = s_1 + s_2 + \cdots + s_n = \delta_{n+4} - (n+3) \quad n \geq 1.$

Para mayor detalle, consúltese [24].

A continuación, presentamos algunas propiedades de las sucesiones $\{a_n\}$ y $\{b_n\}$, dadas en (2.2), en el siguiente lema.

Lema 2.1. *Sean las sucesiones $\{a_n\}$ y $\{b_n\}$ definidas en (2.2) y $\lambda \in [0,1)$ fijo. Si*

$$a_{-1} = \frac{\eta}{\alpha + \eta} < \frac{3 - \sqrt{5}}{2} \qquad y \qquad b_{-1} = \frac{K\beta\alpha^2}{\alpha + \eta} < \frac{a_{-1}(1 - a_{-1})^2}{1 + \lambda(2a_{-1} - 1)}, \tag{2.5}$$

entonces

(i) $\{a_n\}$ y $\{b_n\}$ son sucesiones decrecientes,

(ii) $\varphi = \dfrac{b_0}{b_{-1}} \in (0,1)$ y $\dfrac{a_0}{1 - a_0} < \varphi$,

(iii) $a_n < \varphi^{\delta_n} a_{n-1}$ *y* $b_n < \varphi^{\delta_{n+1}} b_{n-1}$*, para todo* $n \geq 1$*,*

(iv) $a_n < \varphi^{s_n} a_0$*, para todo* $n \geq 1$*.*

Demostración. Para probar (i), procedemos por inducción. De las hipótesis se tiene que $a_0 < a_{-1}$ y $b_0 < b_{-1}$. Si se verifican $a_{j-1} > a_j$ y $b_{j-1} > b_j$, para $j = 0, 1, \ldots, n$, entonces

$$a_{n+1} < f(a_{n-1})g(a_{n-1})b_{n-1} = a_n \qquad \text{y} \qquad b_{n+1} < f(a_{n-1})^2 a_{n-1} b_{n-1} = b_n.$$

Luego $\{a_n\}$ y $\{b_n\}$ son decrecientes.

El apartado (ii) es inmediato por hipótesis.

Para probar (iii), aplicamos inducción. De $a_0 < a_{-1}$ y $b_0 < \varphi b_{-1}$ se deduce que

$$a_1 < f(a_{n-1})g(a_{n-1})\varphi b_{-1} = \varphi a_0 \qquad \text{y} \qquad b_1 < f(a_{n-1})^2 g(a_{-1})\varphi b_{-1} = \varphi b_0.$$

Si suponemos $b_j < \varphi \delta_{j+1} b_{j-1}$, para $j = 1, 2. \ldots, n$, entonces

$$a_{n+1} < f(a_{n-1})g(a_{n-1})\varphi^{\delta_{n+1}} b_{n-1} = \varphi^{\delta_{n+2}} a_n,$$

$$b_{n+1} < f(a_{n-1})^2(\varphi^n g(a_{n-1}))\varphi^{\delta_{n+1}} b_{n-1} < \varphi^{\delta_{n+1}+\delta_n} b_n = \varphi^{\delta_{n+2}} a_n.$$

Finalmente, (iv) es consecuencia de (iii). ■

A continuación, probamos el teorema de convergencia semilocal.

Teorema 2.2. *Sean* X *e* Y *dos espacios de Banach,* $F : \Omega \subset X \to Y$ *un operador definido en un conjunto abierto convexo no vacío* Ω *y* $\lambda \in [0, 1)$*. Suponemos que se cumplen las condiciones (C1)–(C4) y (2.5). Si* $B(x_0, r_0) \subset \Omega$*, donde* $r_0 = \dfrac{1 - a_0}{1 - 2a_0}\eta$*, entonces los métodos tipo secante convergen a una solución* x^* *de* $F(x) = 0$*. Además,* $x_n, x^* \in \overline{B(x_0, r_0)}$ *y* x^* *es única en* $B(x_0, \tau) \cap \Omega$*, donde* $\tau = \frac{1}{K\beta} - r_0 - (1 - \lambda)\alpha$*.*

Demostración. Tenemos que $a_0 < 1$ y $a_n < 1$ para todo $n \geq 1$. Veamos que $\{x_n\}$ es una sucesión de Cauchy y que $x_n \in B(x_0, r_0)$, para todo $n \geq 0$. Ahora, para $m \geq 1$, vemos que

$$\begin{aligned}
\|x_{n+m} - x_n\| &\leq \|x_{n+m} - x_{n+m-1}\| + \|x_{n+m-1} - x_{n+m-2}\| + \cdots + \|x_{n+1} - x_n\| \\
&\leq f(a_{n+m-2})a_{n+m-2} \cdots f(a_{n+1})a_{n+1} f(a_n)a_n \|x_{n+1} - x_n\| \\
&\quad + f(a_{n+m-3})a_{n+m-3} \cdots f(a_{n+1})a_{n+1} f(a_n)a_n \|x_{n+1} - x_n\| \\
&\quad + \cdots + f(a_n)a_n \|x_{n+1} - x_n\| + \|x_{n+1} - x_n\| \\
&= \left(1 + \sum_{i=n}^{n+m-2} \left(\prod_{j=n}^{i} f(a_j)a_j\right)\right) \|x_{n+1} - x_n\|. \qquad (2.6)
\end{aligned}$$

Como $\{a_j\}$ es decreciente y f es creciente, por el lema 2.1 y (2.6), para $n \geq 2$, se tiene

$$\|x_{n+m} - x_n\| = \prod_{j=n}^{n+m-2} \varphi^{s_j} \Delta^n \left[\Delta^{m-1} + \Delta^{m-2} + \cdots + 1\right] \|x_1 - x_0\|,$$

donde $\Delta = \frac{a_0}{1-a_0} < 1$ y, por tanto, $a_0 < \frac{1}{2}$. Entonces,

$$\|x_{n+m} - x_n\| = \left(\varphi^{s_1+s_2+\cdots+s_{n-1}}\right) \frac{\Delta^n(1-\Delta^m)}{1-\Delta}\|x_1 - x_0\|. \tag{2.7}$$

En (2.6), si $n = 1$, se tiene

$$\|x_{m+1} - x_1\| < \frac{\Delta(1-\Delta^m)}{1-\Delta}\|x_1 - x_0\|, \tag{2.8}$$

y, si $n = 0$,

$$\|x_m - x_0\| < \frac{1-\Delta^m}{1-\Delta} < \frac{\eta}{1-\Delta} = r_0. \tag{2.9}$$

Por tanto, $x_n \in B(x_0, r_0)$, para todo $n \geq 1$, la sucesión $\{x_n\}$ está bien definida y es una sucesión de Cauchy. Luego, $\lim\limits_{n\to\infty} x_n = x^* \in \overline{B(x_0, r_0)}$.

Además, por (C4), existe F' y cumple $\|F'(x)-F'(y)\| = \|[x, x; F]-[y, y; F]\| \leq 2K\|x-y\|$, de manera que

$$F(x_n) = (F'(x_{n-1}) - A_{n-1})\,(x_n - x_{n-1}) + \int_0^1 (F'(x_{n-1} + t(x_n - x_{n-1})) - F'(x_{n-1})(x_n - x_{n-1})\,dt,$$

$$\|F(x_n)\| \leq K\left((1-\lambda)\|x_{n-1} - x_{n-2}\| + \|x_n - x_{n-1}\|\right)\|x_n - x_{n-1}\|,$$

$\lim\limits_{n\to\infty} \|F(x_n)\| = 0$ y, por la continuidad de F, vemos que x^* es solución de $F(x) = 0$, puesto que $\lim\limits_{n\to\infty} \|F(x_n)\| = \|F(x^*)\| = 0$.

A continuación, probamos la unicidad de x^*. Sea z^* otra solución distinta de $F(x) = 0$ en $B(x_0, \tau) \cap \Omega$, donde $\tau = \frac{1}{K\beta} - r_0 - (1-\lambda)\alpha$. Si consideramos

$$F(z^*) - F(x^*) = \int_{x^*}^{z^*} F'(x)dx = \int_0^1 F'\left(x^* + t(z^* - x^*)\right)(z^* - x^*)dt = 0,$$

y el operador $J = \displaystyle\int_0^1 F'\left(x^* + t(z^* - x^*)\right)dt$, entonces

$$\begin{aligned}
\|I - A_0^{-1}J\| &\leq \|A_0^{-1}\|\,\|A_0 - J\| \\
&\leq \|A_0^{-1}\| \int_0^1 \|F'\left(x^* + t(z^* - x^*)\right) - A_0\|\ dt \\
&= \|A_0^{-1}\| \int_0^1 \|F'\left(x^* + t(z^* - x^*)\right) - F'(x_0) + F'(x_0) - A_0\|\ dt \\
&\leq \beta\left(\int_0^1 2K\|x^* + t(z^* - x^*) - x_0\|\,dt + \|F'(x_0) - A_0\|\right) \\
&\leq \beta\left(\int_0^1 2K((1-t)\|x^* - x_0\| + t\,\|z^* - x_0\|)\,dt + K(1-\lambda)\alpha\right) \\
&= K\beta\left((1-\lambda)\alpha + \|x^* - x_0\| + \|z^* - x_0\|\right) \\
&< K\beta\left((1-\lambda)\alpha + r_0 + \tau\right) \\
&= 1,
\end{aligned}$$

de manera que J es inversible y, por tanto, $z^* = x^*$. ■

2.1.2. Accesibilidad

Sabemos que los puntos de salida de un método iterativo tienen asociados los parámetros dados en las condiciones iniciales. Para representar la región de accesibilidad del método iterativo, coloreamos los puntos cuyos parámetros asociados verifican las condiciones de convergencia y, en otro caso, no los coloreamos. La región de accesibilidad asociada a una solución de $F(x) = 0$ nos indica entonces el dominio de puntos de salida a partir de los cuales tenemos asegurada la convergencia del método iterativo que se aplica; es decir, el conjunto de puntos de salida que satisfacen las condiciones de convergencia impuestas al método iterativo.

A continuación, vemos cuál es la región de accesibilidad asociada a la raíz $z = 1$ de la ecuación compleja $F(z) = z^3 - 1 = 0$ cuando se aproxima mediante los métodos tipo secante. Considerando el cuadrado $[0.9, 1.3] \times [-0.2, 0.2]$ como dominio complejo, que únicamente contiene la raíz $z = 1$, obtenemos $K = 6|1.3 + 0.2i| = 3.9458\ldots$ Tomando $z_{-1} = z_0 - d$, representamos la región de accesibilidad coloreando los puntos z_0 que verifiquen las condiciones de convergencia dadas en (2.5). Así, fijado $\alpha = |z_0 - z_{-1}| = |d|$, en las figuras 2.1–2.4 se muestran las regiones de accesibilidad para cuatro métodos tipo secante: $\lambda = 0$ (región verde), $\lambda = \frac{1}{4}$ (región rosa), $\lambda = \frac{1}{2}$ (región amarilla) y $\lambda = \frac{3}{4}$ (región morada).

Se observa entonces que las regiones de accesibilidad de los métodos tipo secante que se han representado gráficamente tienen el problema de que no se puede garantizar la convergencia para ciertos valores de λ, aún estando cerca de la raíz o en la misma raíz. Vemos que aparece una zona hueca en la región de accesibilidad que contiene a la propia raíz. Evidentemente, esta restricción es consecuencia de que la distancia entre los puntos de salida (o, equivalentemente, del valor de α) no es suficientemente pequeña como para poder garantizar la convergencia en estas situaciones concretas.

Por otra parte, si se consideran valores de α más pequeños, se puede ver en las figuras 2.5–2.8 que, aunque se va reduciendo la zona hueca, también se reduce la región de accesibilidad. De hecho, se puede comprobar que para $\alpha = 1/64$ ya no existe región de accesibilidad, como consecuencia de que las condiciones (2.10) no se cumplen.

A continuación, estudiando el dominio de parámetros asociado al teorema 2.2, vamos a detallar cuáles son los problemas de accesibilidad de los métodos tipo secante. Para representarlo gráficamente, se colorean en un plano xy los valores de los parámetros correspondientes a los puntos de salida que verifican las condiciones que se imponen en el teorema 2.2. Observamos que las condiciones exigidas a los puntos de salida, (C1)–(C3), introducen los parámetros α, β y η, y la condición exigida al operador F, (C4), introduce el parámetro fijo K. En primer lugar, expresamos las dos condiciones dadas en (2.5) de forma explícita a partir de los valores que consideramos para construir el dominio de parámetros: $K\beta\eta$ y $K\beta\alpha$. Así, podemos escribir (2.5) de la siguiente forma:

$$\frac{K\beta\eta}{K\beta\alpha + K\beta\eta} < \frac{3-\sqrt{5}}{2} \qquad \text{y} \qquad 1 < \frac{K\beta\eta}{(K\beta\alpha + K\beta\eta)(K\beta\alpha + K\beta\eta + \lambda(K\beta\eta - K\beta\alpha)}. \tag{2.10}$$

Ahora, considerando que representamos en el eje x de abscisas los valores $K\beta\eta$ y en el y de ordenadas los valores $K\beta\alpha$, dibujamos en la figura 2.9 los dominios de parámetros que están sujetos a las condiciones dadas en (2.10) para cuatro métodos tipo secante: $\lambda = 0$ (región verde), $\lambda = 0.25$ (región rosa), $\lambda = 0.5$ (región amarilla) y $\lambda = 0.75$ (región morada). Notamos que las regiones están superpuestas.

A partir de la figura 2.9 observamos dos situaciones que destacamos a continuación. En primer lugar, la elección de adecuados puntos de salida x_{-1} y x_0 para los métodos tipo secante

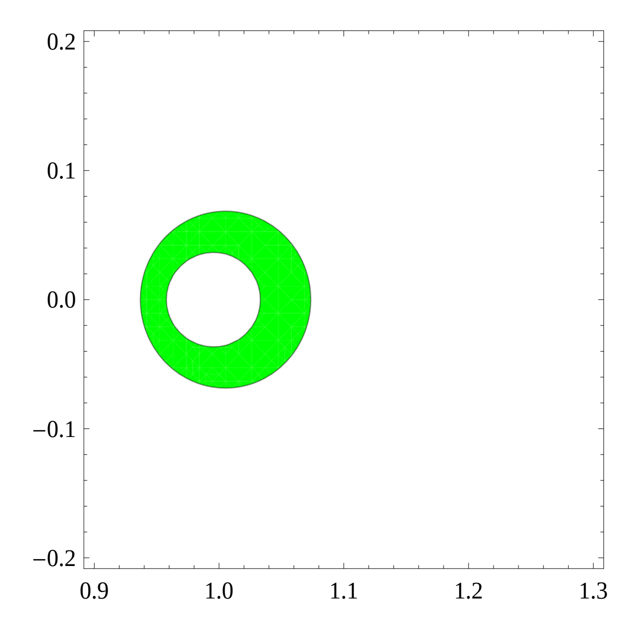


Figura 2.1: Región de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método tipo secante correspondiente a $\lambda = 0$ y tomando $\alpha = \frac{1}{10}$.

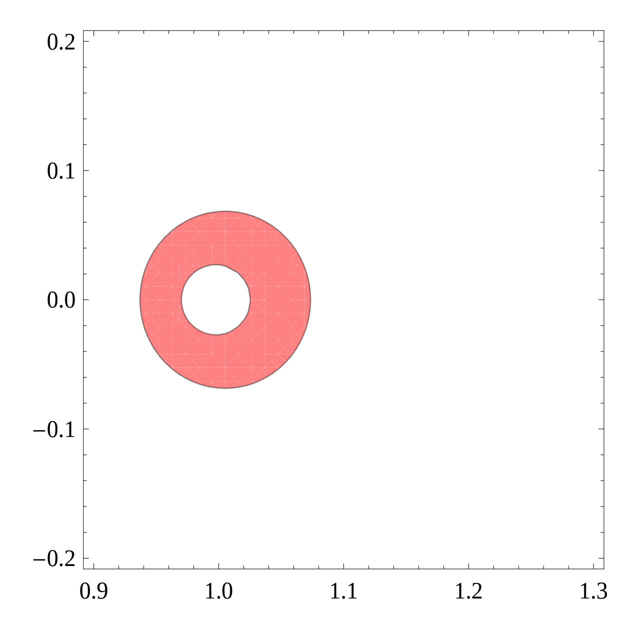


Figura 2.2: Región de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método tipo secante correspondiente a $\lambda = \frac{1}{4}$ y tomando $\alpha = \frac{1}{10}$.

es muy restrictiva, puesto que el dominio de parámetros es muy reducido. En segundo lugar, si consideramos un valor fijo de $K\beta\alpha$ perteneciente al dominio de parámetros, observamos que los posibles valores que se pueden considerar de $K\beta\eta$, para que los puntos de salida pertenezcan al dominio de parámetros, tienen una cota superior y una cota inferior para la cantidad $K\beta\eta$. Esto hace que, incluso tomando la propia raíz como punto de salida ($K\beta\eta = 0$), no obtengamos puntos iniciales que verifiquen las condiciones del teorema 2.2, lo que resulta evidente a partir de la segunda condición de (2.10), puesto que obviamente nunca se verifica la desigualdad para $\alpha > 0$ y $\eta = 0$ (es decir, $x_0 = x^*$).

2.2. Método predictor: el método simplificado de la secante

Una vez descritos los problemas de accesibilidad que se deducen del teorema 2.2 para los métodos tipo secante, introducimos ahora el método simplificado de la secante (2.1) y analizamos su convergencia semilocal. A partir de este estudio, veremos que este método tiene mejor accesibilidad que los método tipo secante, lo que posteriormente utilizamos para considerarlo como método predictor del método híbrido (predictor-corrector) que construimos.

2.2.1. Convergencia semilocal

Para probar la convergencia semilocal del método simplificado de la secante (2.1), suponemos que se cumplen las siguientes condiciones iniciales:

(H1) $\|z_0 - z_{-1}\| = \alpha_0 \neq 0$, con $z_{-1}, z_0 \in \Omega$,

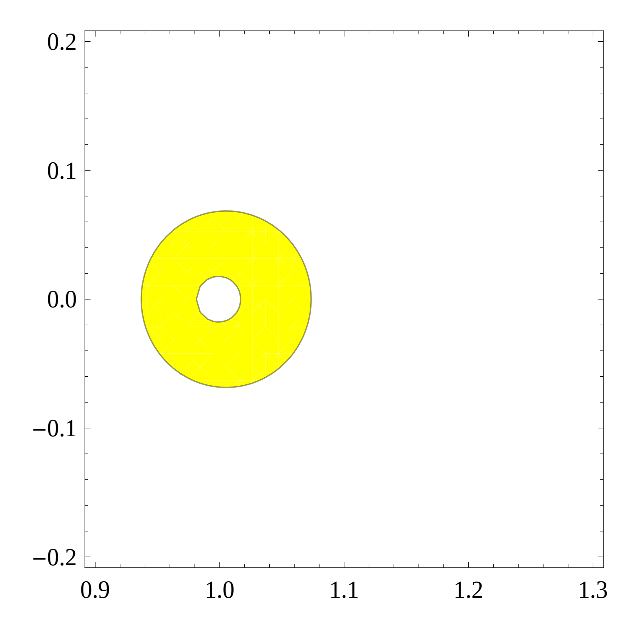


Figura 2.3: Región de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método tipo secante correspondiente a $\lambda = \frac{1}{2}$ y tomando $\alpha = \frac{1}{10}$.

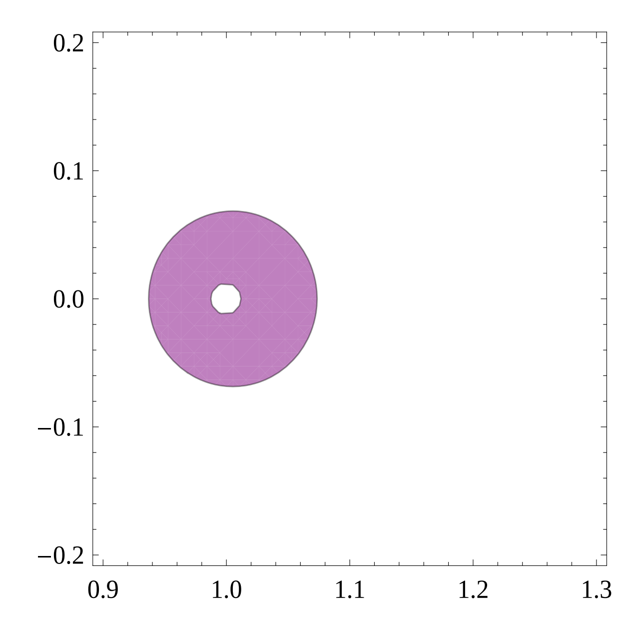


Figura 2.4: Región de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método tipo secante correspondiente a $\lambda = \frac{3}{4}$ y tomando $\alpha = \frac{1}{10}$.

(H2) existe $L_0^{-1} = [z_{-1}, z_0; F]^{-1} \in \mathcal{L}(Y, X)$, para $z_{-1}, z_0 \in \Omega$, y es tal que $\|L_0^{-1}\| \leq \gamma$,

(H3) $\|L_0^{-1} F(z_0)\| \leq \varepsilon$,

(H4) $\|[z, y; F] - [u, v; F]\| \leq K(\|z - u\| + \|y - v\|)$, $K \geq 0$, $z, y, u, v \in \Omega$, $z \neq y$, $u \neq v$.

Antes de probar la convergencia semilocal del método simplificado de la secante (2.1) bajo las condiciones (H1)–(H4), probamos el siguiente lema técnico.

Lema 2.3. *A partir de (H1)–(H4), si*

$$K\gamma\alpha_0 < 1 \qquad y \qquad K\gamma\varepsilon < \frac{\sqrt{2(1 + (K\gamma\alpha_0)^2)} - (1 + K\gamma\alpha_0)}{2}, \tag{2.11}$$

entonces la ecuación

$$2K\gamma t^2 - (1 - K\gamma\alpha_0 + 2K\gamma\varepsilon)t + (1 + K\gamma\varepsilon)\varepsilon = 0 \tag{2.12}$$

tiene dos raíces reales positivas. Si denotamos por R la menor de ellas, se tiene que $K\gamma(2R + \alpha_0) < 1$ y $R > \varepsilon$.

Demostración. Como el discriminante de la ecuación (2.12) está dado por

$$\begin{aligned} \Delta &= (1 - K\gamma\alpha_0 + 2K\gamma\varepsilon)^2 - 8K\gamma\varepsilon(1 + K\gamma\varepsilon) \\ &= \left((1 - K\gamma\alpha_0 + 2K\gamma\varepsilon) + \sqrt{8K\gamma\varepsilon(1 + K\gamma\varepsilon)}\right) \\ &\quad \times \left((1 - K\gamma\alpha_0 + 2K\gamma\varepsilon) - \sqrt{8K\gamma\varepsilon(1 + K\gamma\varepsilon)}\right), \end{aligned}$$

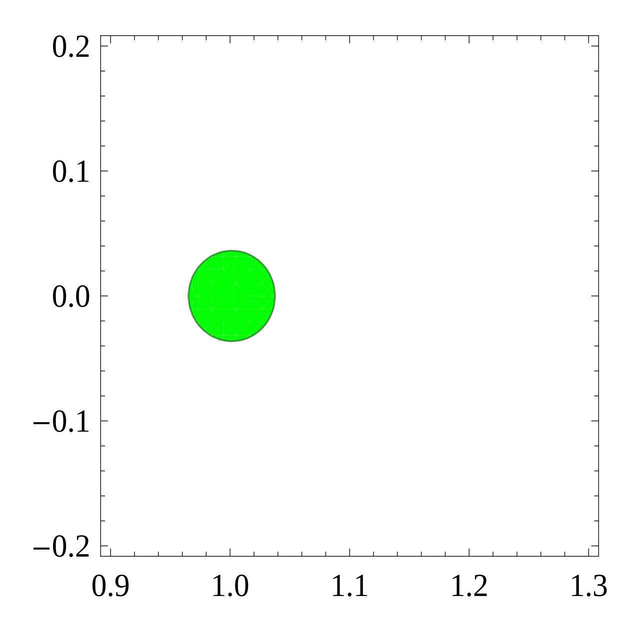


Figura 2.5: Región de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método tipo secante correspondiente a $\lambda = 0$ y tomando $\alpha = \frac{1}{16}$.

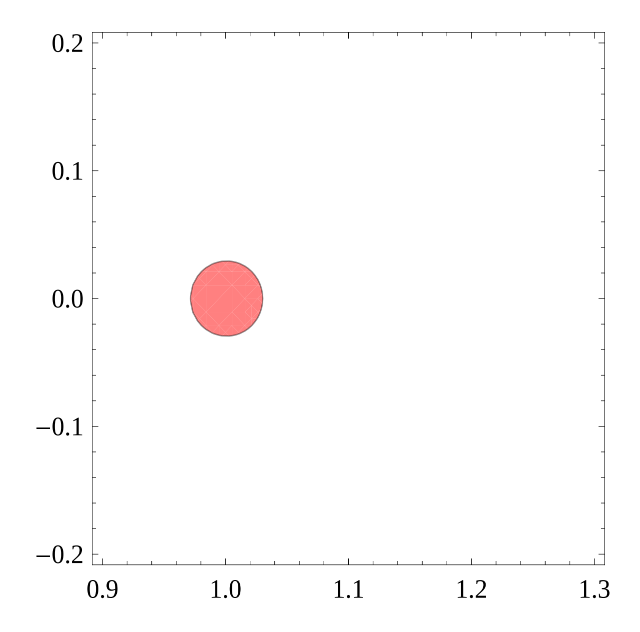


Figura 2.6: Región de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método tipo secante correspondiente a $\lambda = \frac{1}{4}$ y tomando $\alpha = \frac{1}{20}$.

la ecuación (2.12) tiene dos raíces reales si y solo si $\Delta > 0$. Analizamos ahora los dos factores de Δ. Por hipótesis, tenemos que $K\gamma\alpha_0 < 1$, de manera que $\Delta > 0$ si

$$1 - K\gamma\alpha_0 + 2K\gamma\varepsilon > \sqrt{8K\gamma\varepsilon(1 + K\gamma\varepsilon)},$$

y operando llegamos a

$$(1 - K\gamma\alpha_0)^2 - 4(1 + K\gamma\alpha_0)K\gamma\varepsilon - 4(K\gamma\varepsilon)^2 > 0,$$

que conduce a la segunda condición de (2.11).

Notemos que

$$1 - K\gamma\alpha_0 + 2K\gamma\varepsilon > \sqrt{8K\gamma\varepsilon(1 + K\gamma\varepsilon)} > 0,$$

de manera que la menor raíz real positiva de (2.12) es:

$$R = \frac{1 - K\gamma\alpha_0 + 2K\gamma\varepsilon - \sqrt{(1 - K\gamma\alpha_0)^2 - 4(1 + K\gamma\alpha_0)K\gamma\varepsilon - 4(K\gamma\varepsilon)^2}}{4K\gamma}. \tag{2.13}$$

Probamos a continuación que $K\gamma(2R+\alpha_0) < 1$. Para ello, probamos previamente que $K\gamma(2\varepsilon+\alpha_0) < 1$. A partir de la segunda condición de (2.11), tenemos

$$K\gamma\alpha_0 + 2K\gamma\varepsilon < \sqrt{2(1 + (K\gamma\alpha_0)^2)} - 1.$$

Además, como

$$\sqrt{2(1 + (K\gamma\alpha_0)^2)} - 1 < 1,$$

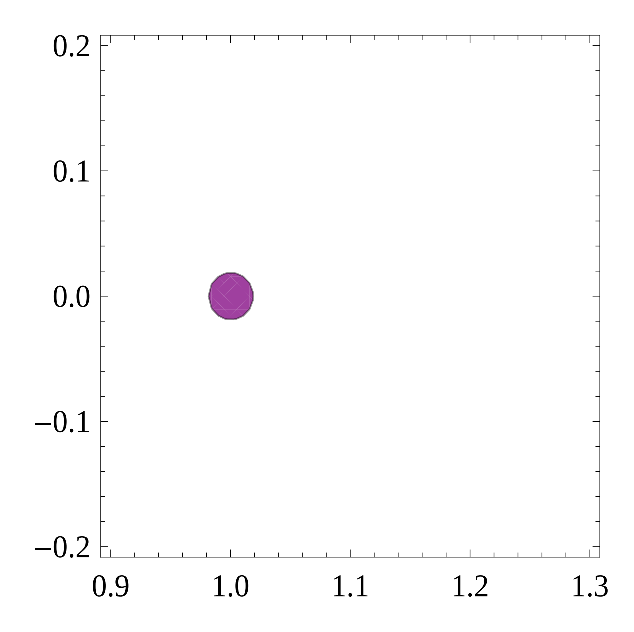


Figura 2.7: Región de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método tipo secante correspondiente a $\lambda = \frac{3}{4}$ y tomando $\alpha = \frac{1}{32}$.

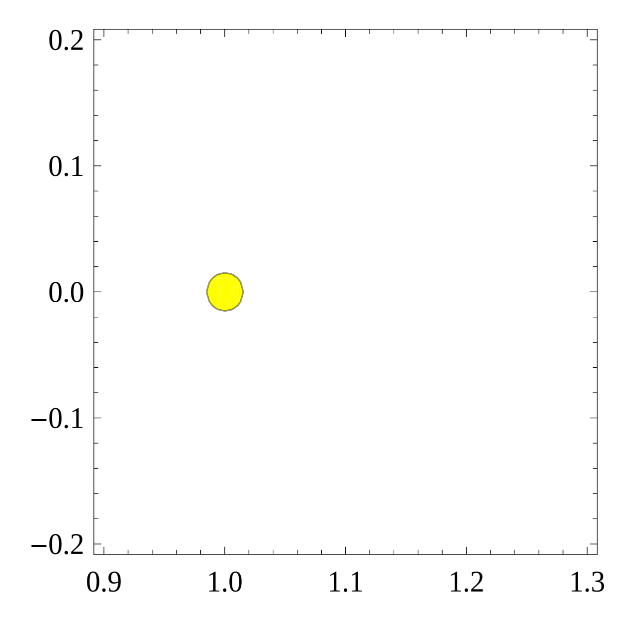


Figura 2.8: Región de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método tipo secante correspondiente a $\lambda = \frac{1}{2}$ y tomando $\alpha = \frac{1}{40}$.

ya que $2(1+(K\gamma\alpha_0)^2) < 4$, por verificarse $K\gamma\alpha_0 < 1$, tenemos $K\gamma\alpha_0+2K\gamma\varepsilon = K\gamma(\alpha_0+2\varepsilon) < 1$. Ahora, observamos que la condición $K\gamma(2R + \alpha_0) < 1$ es equivalente a la condición

$$1 + K\gamma\alpha_0 + 2K\gamma\varepsilon - \sqrt{(1 - K\gamma\alpha_0 + 2K\gamma\varepsilon)^2 - 8K\gamma\varepsilon(1 + K\gamma\varepsilon)} < 2,$$

que se satisface trivialmente porque $K\gamma(\alpha_0 + 2\varepsilon) < 1$. La desigualdad $R > \varepsilon$ resulta fácil de probar a partir de (2.13). ■

Observemos que en el resultado anterior también podemos considerar la existencia de una raíz doble. Para ello, basta considerar las desigualdades no estrictas.

A continuación, damos un lema técnico para la sucesión $\{z_n\}$ dada por el método simplificado de la secante (2.1).

Lema 2.4. *Sea $\{z_n\}$ la sucesión dada por el método simplificado de la secante (2.1). Supongamos (H1)–(H4). Si $z_{n-1} \neq z_n$ con $z_{n-1}, z_n \in \Omega$, entonces*

(i) $F(z_n) = (L_n - L_0)(z_n - z_{n-1})$, *donde* $L_0 = [z_{-1}, z_0; F]$ *y* $L_n = [z_{n-1}, z_n; F]$,

(ii) $\|z_{n+1} - z_n\| \leq K\gamma(\|z_n - z_0\| + \|z_{n-1} - z_0\| + \|z_0 - z_{-1}\|)\|z_n - z_{n-1}\|$.

Demostración. A partir del algoritmo del método simplificado de la secante (2.1), tenemos $F(z_{n-1}) + L_0(z_n - z_{n-1}) = 0$, de manera que

$$F(z_n) = F(z_n) - F(z_{n-1}) - L_0(z_n - z_{n-1}) = (L_n - L_0)(z_n - z_{n-1}).$$

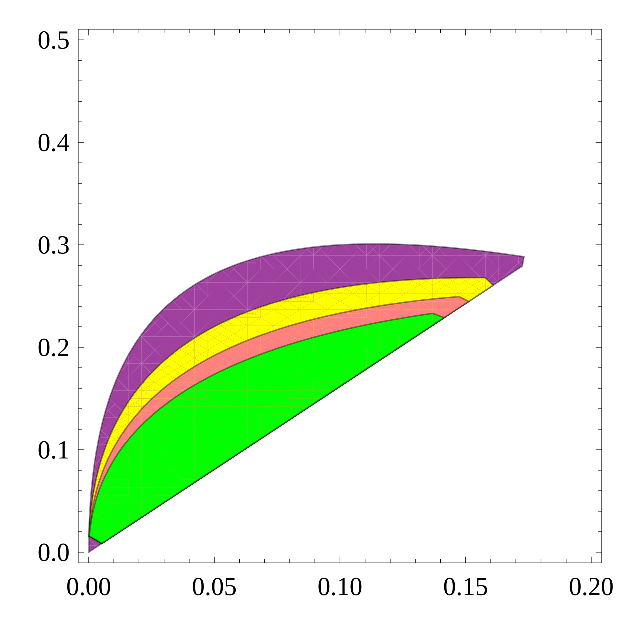


Figura 2.9: Dominios de parámetros de los métodos tipo secante asociados al teorema 2.2 cuando $\lambda = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ (regiones verde, rosa, amarilla y morada, respectivamente).

Por otra parte,

$$\begin{aligned} \|z_{n+1} - z_n\| &\leq \|L_0^{-1}\|\|F(z_n)\| \\ &\leq \|L_0^{-1}\|\|L_n - L_0\|\|z_n - z_{n-1}\| \\ &\leq K\gamma(\|z_n - z_0\| + \|z_{n-1} - z_{-1}\|)\|z_n - z_{n-1}\| \\ &\leq K\gamma(\|z_n - z_0\| + \|z_{n-1} - z_0\| + \|z_0 - z_{-1}\|)\|z_n - z_{n-1}\|. \quad \blacksquare \end{aligned}$$

A continuación, presentamos el siguiente resultado de convergencia semilocal para el método simplificado de la secante (2.1).

Teorema 2.5. *Sean X e Y dos espacios de Banach y $F : \Omega \subset X \to Y$ un operador definido en un conjunto abierto convexo no vacío Ω. Suponemos que se cumplen las condiciones (H1)–(H4) y (2.11). Si $B(z_0, R) \subset \Omega$, con R dado en (2.13), entonces la sucesión $\{z_n\}$ dada por el método simplificado de la secante (2.1) está bien definida y converge a una solución z^* de la ecuación $F(x) = 0$. Además $z_n, z^* \in \overline{B(z_0, R)}$ y z^* es única en $B(z_0, r) \cap \Omega$, donde $r = \frac{1}{K\gamma} - R - \alpha_0$.*

Demostración. Comenzamos probando que la sucesión $\{z_n\}$ está bien definida; es decir, $z_n \in B(z_0, R) \subset \Omega$, para todo $n \in \mathbb{N}$. Lo haremos por inducción matemática sobre n. En primer lugar, consideramos $z_1 = z_0 - L_0^{-1}F(z_0)$, donde $L_0 = [z_{-1}, z_0; F]$. En este caso, por el lema 2.3, tenemos

$$\|z_1 - z_0\| = \|L_0^{-1}F(z_0)\| \leq \varepsilon < R.$$

Luego, $z_1 \in B(z_0, R) \subset \Omega$ y podemos definir $z_2 = z_1 - L_0^{-1}F(z_1)$.

A continuación, por el lema 2.4, tenemos

$$\|z_2 - z_1\| \leq K\gamma(\|z_1 - z_0\| + \|z_0 - z_{-1}\|)\|z_1 - z_0\| \leq K\gamma(\varepsilon + \alpha_0)\varepsilon.$$

Por tanto, de (2.12), (2.13) y $K\gamma(2R+\alpha_0)<1$, se sigue

$$\begin{aligned}
\|z_2-z_0\| &\leq \|z_2-z_1\|+\|z_1-z_0\| \\
&\leq (1+K\gamma(\varepsilon+\alpha_0))\varepsilon \\
&\leq \left(1+\frac{K\gamma(\varepsilon+\alpha_0)}{1-K\gamma(2R+\alpha_0)}\right)\varepsilon \\
&= R.
\end{aligned}$$

Luego, $z_2 \in B(z_0,R)\subset\Omega$ y podemos definir $z_3 = z_2 - L_0^{-1}F(z_2)$.

Utilizando inducción matemática sobre n, suponemos que $z_j \in B(z_0,R)\subset\Omega$, para $j = 2,3,\dots,n$,

$$\|z_n - z_{n-1}\| < K\gamma(2R+\alpha_0)\|z_{n-1}-z_{n-2}\|,$$

$$\|z_n - z_0\| \leq \left(1+K\gamma(\varepsilon+\alpha_0)\sum_{i=0}^{n-2}(K\gamma(2R+\alpha_0))^i\right)\varepsilon < R.$$

Entonces, $z_{n+1} = z_n - L_0^{-1}F(z_n)$ está bien definido. Además,

$$\|z_{n+1}-z_n\| \leq K\gamma(\|z_n-z_0\|+\|z_{n-1}-z_0\|+\|z_0-z_{-1}\|)\|z_n-z_{n-1}\| < K\gamma(2R+\alpha_0)\|z_n-z_{n-1}\|$$

$$\begin{aligned}
\|z_{n+1}-z_0\| &\leq \|z_{n+1}-z_n\|+\|z_n-z_0\| \\
&\leq K\gamma(2R+\alpha_0)\|z_n-z_{n-1}\|+\|z_n-z_0\| \\
&\leq (K\gamma(2R+\alpha_0))^{n-1}\|z_2-z_1\|+\|z_n-z_0\| \\
&\leq \left(K\gamma(\varepsilon+\alpha_0)(K\gamma(2R+\alpha_0))^{n-1}+K\gamma(\varepsilon+\alpha_0)\sum_{j=0}^{n-2}(K\gamma(2R+\alpha_0))^j+1\right)\varepsilon \\
&< \left(1+\frac{K\gamma(\varepsilon+\alpha_0)}{1-K\gamma(2R+\alpha_0)}\right)\varepsilon \\
&= R
\end{aligned}$$

y la sucesión $\{z_n\}$ está entonces bien definida.

Por otra parte, resulta evidente que

$$\|z_{n+1}-z_n\| \leq K\gamma(\varepsilon+\alpha_0)(K\gamma(2R+\alpha_0))^{n-1},$$

y, como $K\gamma(2R+\alpha_0)<1$, se sigue que $\{z_n\}$ es una sucesión de Cauchy, y por tanto convergente a un punto $z^* \in \overline{B(z_0,R)}$. Veamos que z^* es solución de la ecuación $F(x)=0$. Como

$$\begin{aligned}
\|F(z_n)\| &\leq \|L_0-L_n\|\|z_n-z_{n-1}\| \\
&\leq K\left(\|z_n-z_0\|+\|z_{n-1}-z_{-1}\|\right)\|z_n-z_{n-1}\| \\
&< K(2R+\alpha_0)\|z_n-z_{n-1}\|,
\end{aligned}$$

por la continuidad del operador F, es fácil ver que $F(z^*)=0$, puesto que $\lim\limits_{n\to\infty}\|F(z_n)\| = \|F(x^*)\| = 0$.

Finalmente, probamos la unicidad de la solución z^* en $B(z_0,r)\cap\Omega$, donde $r=\frac{1}{K\gamma}-R-\alpha_0$. Suponemos entonces que tenemos otra solución distinta $y^* \in B(z_0,r)\cap\Omega$ de $F(x)=0$. Consideramos

$$F(y^*)-F(z^*) = \int_{z^*}^{y^*} F'(u)\,du = \int_0^1 F'(z^*+t(y^*-z^*))(y^*-z^*)\,dt = 0,$$

y el operador $J = \int_0^1 F'(z^* + t(y^* - z^*))\, dt$. Teniendo en cuenta

$$\begin{aligned}
\|I - L_0^{-1}J\| &\leq \|L_0^{-1}\|\|L_0 - J\| \\
&\leq \|L_0^{-1}\| \int_0^1 \|F'(z^* + t(y^* - z^*)) - L_0\|\, dt \\
&\leq \|L_0^{-1}\| \int_0^1 \left(\|F'(z^* + t(y^* - z^*)) - F'(z_0)\| + \|F'(z_0) - L_0\|\right) dt \\
&\leq \gamma \int_0^1 2K\|z^* + t(y^* - z^*) - z_0\|\, dt + \gamma\|F'(z_0) - L_0\| \\
&\leq \gamma \int_0^1 2K\left((1-t)\|z^* - z_0\| + t\|y^* - z_0\|\right)\, dt + K\gamma\|z_0 - z_{-1}\| \\
&< K\gamma(R + r + \alpha_0) \\
&= 1
\end{aligned}$$

y el lema de Banach (lema 1.22), se sigue que J es inversible y, por tanto, $y^* = z^*$. Notemos que $r > R > 0$, puesto que $K\gamma(2R + \alpha_0) < 1$, por cumplirse (2.11) (véase el lema 2.3). ■

2.2.2. Accesibilidad

Si consideramos de nuevo la ecuación compleja $F(z) = z^3 - 1 = 0$ y observamos las regiones de accesibilidad asociadas a la raíz $z = 1$ para el método simplificado de la secante (2.1) (región roja) y los métodos tipo secante para distintos valores de λ ($\lambda = 0$ (región verde), $\lambda = \frac{1}{4}$ (región rosa), $\lambda = \frac{1}{2}$ (región amarilla) y $\lambda = \frac{3}{4}$ (región morada)) en las condiciones indicadas anteriormente, vemos claramente la mejora que se consigue tomando distintos valores del parámetro α_0 en las figuras 2.10–2.17. Notamos que las regiones están superpuestas.

Nuestro siguiente objetivo es comparar con mayor exactitud los dominios de parámetros de los métodos tipo secante con el del simplificado de la secante (2.1). Para ello, tenemos que representar los mismos valores en los ejes de los planos donde representamos gráficamente los dominios de los parámetros. Para ello, tenemos que escribir β en función de γ, de manera que podemos representar los valores de los inversos de las mismas diferencias divididas. Así, si L_0^{-1} existe y $\|L_0^{-1}\| \leq \varepsilon$, tenemos

$$\|I - L_0^{-1}A_0\| \leq \|L_0^{-1}\|\,\|L_0 - A_0\| \leq \gamma K\|y_0 - x_{-1}\| \leq \gamma K\lambda\alpha$$

y, siempre que $\gamma K\lambda\alpha < 1$,

$$\|A_0^{-1}\| \leq \frac{\gamma}{1 - \gamma K\lambda\alpha} = \beta.$$

Como se puede observar en la figura 2.18, el dominio de parámetros del método simplificado de la secante (2.1) resuelve el problema que tenían los métodos tipo secante cuando los valores de α_0 o ε son pequeños. Fijado un valor de $K\gamma\alpha_0$ perteneciente al dominio de parámetros, observamos que el valor de $K\gamma\varepsilon$ solo está acotado superiormente (no inferiormente). Por ello, en las regiones de accesibilidad del método simplificado de la secante (2.1) no aparecen zonas huecas conteniendo a la raíz, tal y como ocurre para los métodos tipo secante. Notamos que las regiones están superpuestas.

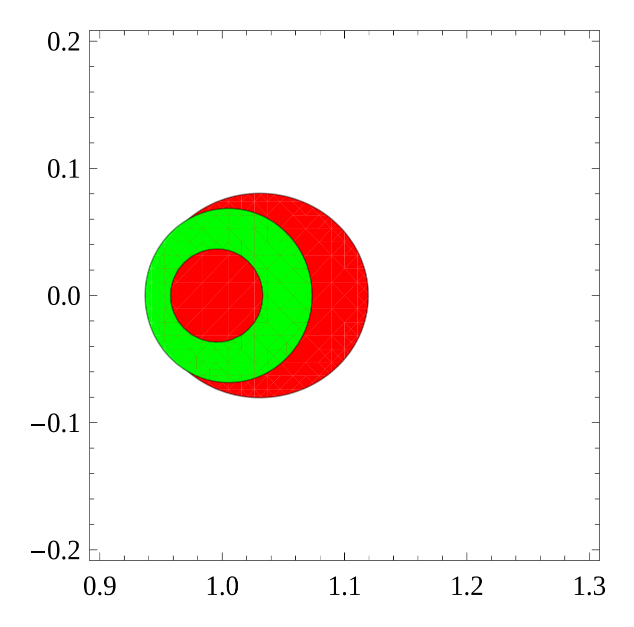


Figura 2.10: Regiones de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método simplificado de la secante (región roja) y el método tipo secante correspondiente a $\lambda = 0$ (región verde) y tomando $\alpha = \frac{1}{8}$.

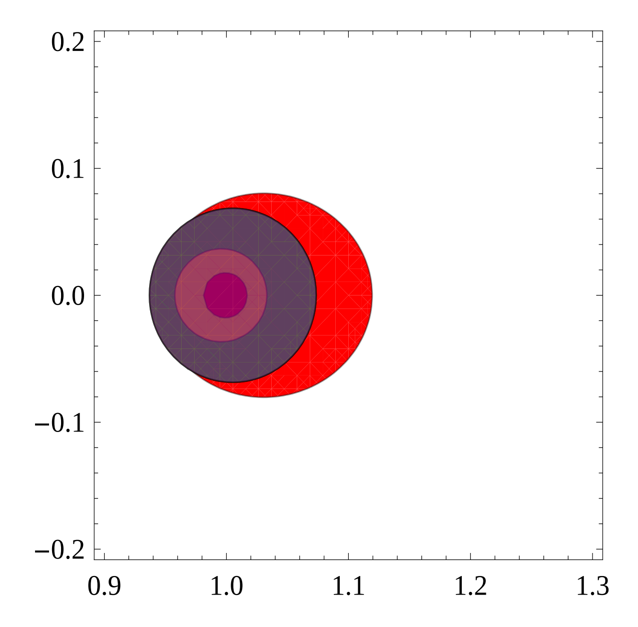


Figura 2.11: Regiones de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método simplificado de la secante (región roja) y el método tipo secante correspondiente a $\lambda = \frac{3}{4}$ (región morada) y tomando $\alpha = \frac{1}{8}$.

2.3. Método iterativo híbrido (predictor-corrector)

A partir de lo visto anteriormente, el objetivo principal es ahora construir, apoyándonos en el método simplificado de la secante (2.1), una modificación de los métodos tipo secante que mejore su dominio de parámetros y su región de accesibilidad.

2.3.1. Construcción del método

Como se observa en la figura 2.18, el dominio de parámetros del método simplificado de la secante (2.1) no tiene los problemas de accesibilidad de los dominios de parámetros de los métodos tipo secante. Trataremos entonces de asegurar que, para una terna inicial $(\alpha_0, \gamma, \varepsilon)$ que satisfaga las condiciones dadas en (2.11) y estar así en el dominio de parámetros del método simplificado de la secante (2.1), podamos obtener una terna (α, β, η) que satisfaga las condiciones dadas en (2.5) después de realizar un cierto número de iteraciones N_0 con el método simplificado de la secante (2.1), de manera que estemos en condiciones de poder garantizar la convergencia de la familia de métodos tipo secante. Cuando esto ocurra, podremos considerar la terna $(\alpha_{N_0}, \beta_{N_0}, \eta_{N_0})$ como terna inicial (α, β, η) de la familia de métodos tipo secante.

Nuestro objetivo inmediato es entonces construir una sencilla modificación de la familia de métodos tipo secante que sea convergente al empezar en los mismos puntos de salida que garantizan la convergencia del método simplificado de la secante (2.1). Así, consideramos el

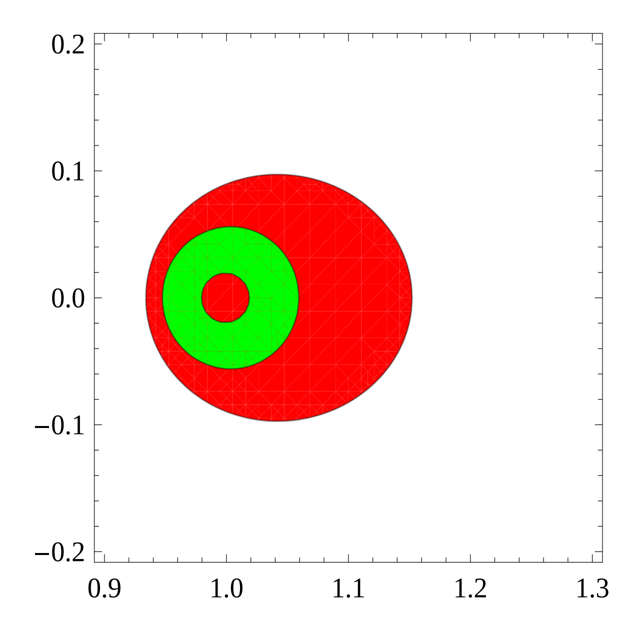


Figura 2.12: Regiones de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método simplificado de la secante (región roja) y el método tipo secante correspondiente a $\lambda = 0$ (región verde) y tomando $\alpha = \frac{1}{10}$.

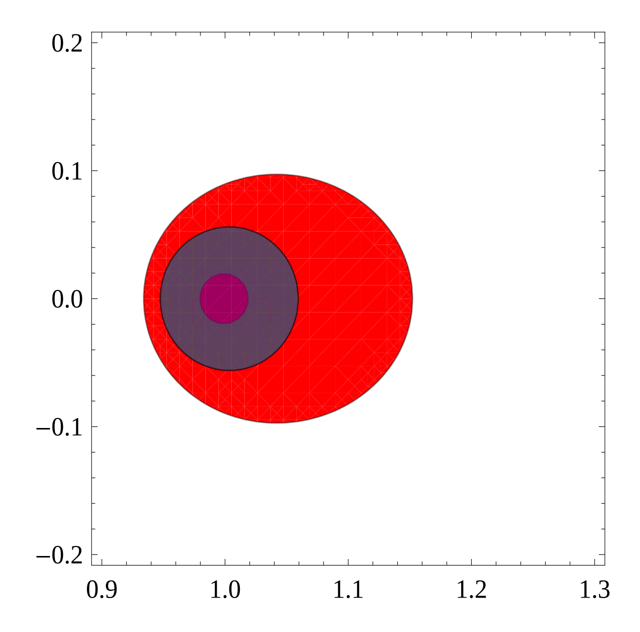


Figura 2.13: Regiones de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método simplificado de la secante (región roja) y el método tipo secante correspondiente a $\lambda = \frac{3}{4}$ (región morada) y tomando $\alpha = \frac{1}{10}$.

método iterativo híbrido (predictor-corrector) dado por

$$\begin{cases} \begin{cases} \text{dados } z_{-1}, z_0 \text{ en } \Omega, \\ z_{i+1} = z_i - [z_0, z_{-1}; F]^{-1} F(z_i), \quad i = 0, 1, \ldots, N_0 - 1, \end{cases} \\ \begin{cases} x_{-1} = z_{N_0-1}, x_0 = z_{N_0}, \\ x_{n+1} = x_n - [\lambda x_n + (1-\lambda) x_{n-1}, x_n; F]^{-1} F(x_n), \quad \lambda \in [0,1), \quad n \geq 0, \end{cases} \end{cases} \tag{2.14}$$

donde z_{-1} y z_0 satisfacen (2.11), mientras que $x_{-1} = z_{N_0-1}$ y $x_0 = z_{N_0}$ satisfacen (2.5). Para que el método híbrido (2.14) sea convergente, nos planteamos dos cuestiones:

1. Localizar z_{-1} y z_0 de manera que el método predictor, el método simplificado de la secante (2.1), sea convergente.

2. Utilizando la convergencia del método predictor, calcular un valor N_0 tal que z_{N_0-1} y z_{N_0} sean considerados puntos iniciales a partir de los cuales la convergencia del método corrector, la familia de métodos tipo secante para un valor fijo de λ, esté garantizada.

Así, utilizamos el método simplificado de la secante (2.1) durante un número finito de pasos N_0 hasta que $z_{N_0-1} = x_{-1}$ y $z_{N_0} = x_0$ cumplan las condiciones dadas en (2.5) y, después, aplicaremos un método de la familia de métodos tipo secante en vez del método simplificado de la secante (2.1). La clave del problema está entonces en garantizar la existencia de N_0.

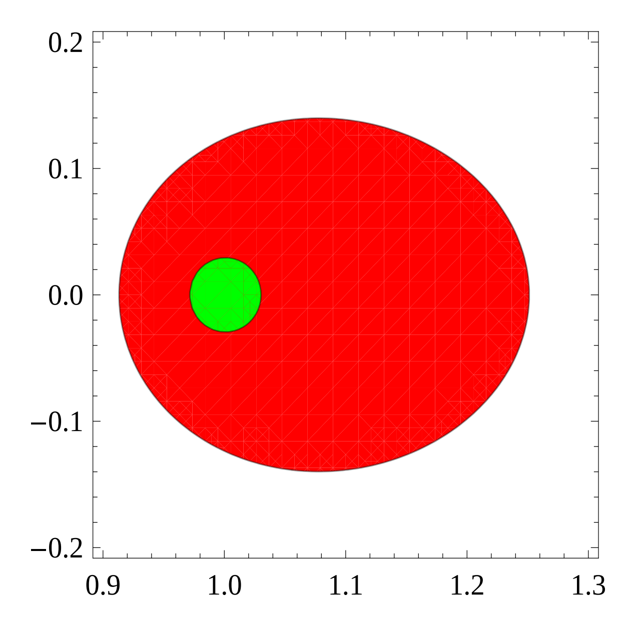


Figura 2.14: Regiones de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método simplificado de la secante (región roja) y el método tipo secante correspondiente a $\lambda = 0$ (región verde) y tomando $\alpha = \frac{1}{20}$.

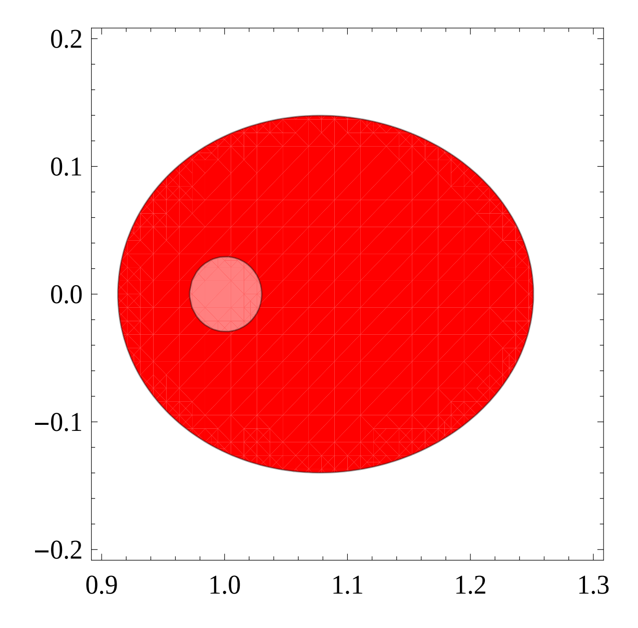


Figura 2.15: Regiones de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método simplificado de la secante (región roja) y el método tipo secante correspondiente a $\lambda = \frac{1}{4}$ (región rosa) y tomando $\alpha = \frac{1}{20}$.

2.3.2. Convergencia semilocal del método

A continuación, vamos a estudiar la convergencia semilocal del método (2.14). A partir del método predictor, consideramos la siguiente situación. Dadas las aproximaciones iniciales z_{-1} y z_0, consideramos la sucesión $\{z_n\}$ definida por el método simplificado de la secante (2.1) y denotamos $M = K\gamma(2R + \alpha_0)$. Para que el método simplificado de la secante (2.1) sea convergente, sabemos que la terna inicial $(\alpha_0, \gamma, \varepsilon)$ debe verificar las condiciones dadas en (2.11). Nos planteamos ahora cómo encontrar N_0 de manera que, a partir de la iteración N_0 del método simplificado de la secante (2.1), podamos considerar la familia de métodos tipo secante con $x_{-1} = z_{N_0-1}$ y $x_0 = z_{N_0}$ cumpliendo (2.5).

En primer lugar, observamos que la definición de la terna inicial $(\alpha_0, \beta_0, \eta_0)$ para la familia de métodos tipo secante es inmediata sin más que tener en cuenta $\beta_0 = \dfrac{\gamma}{1-M}$ y $\eta_0 = \varepsilon$, puesto que

$$\|I - L_0^{-1}A_0\| \leq \|L_0^{-1}\| \, \|L_0 - A_0\| \leq \gamma K \lambda \alpha_0 < \gamma K \alpha_0 < M < 1.$$

A continuación, vamos iterando para definir las ternas $(\alpha_n, \beta_n, \eta_n)$ asociadas a cada z_n. Para ello, procedemos de la siguiente forma.

Primer paso del método predictor: definición de la terna $(\alpha_1, \beta_1, \eta_1)$.

Como

$$\|z_1 - z_0\| = \|L_0^{-1}F(z_0)\| \leq \varepsilon = \eta_0 = \alpha_1,$$

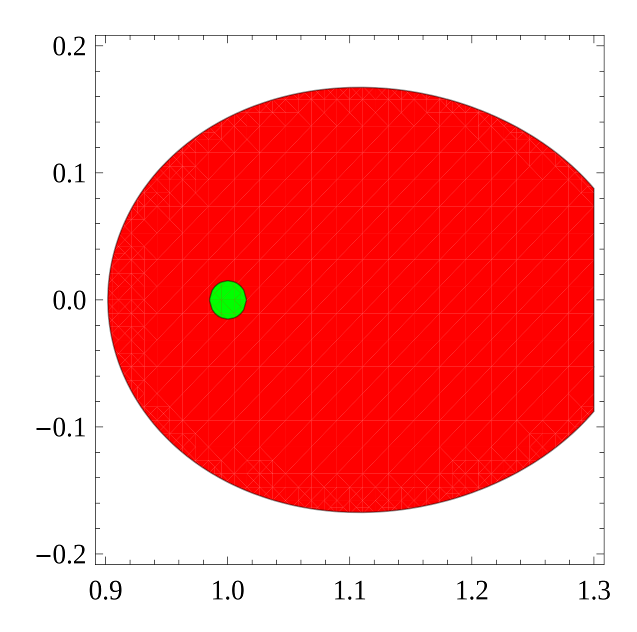


Figura 2.16: Regiones de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método simplificado de la secante (región roja) y el método tipo secante correspondiente a $\lambda = 0$ (región verde) y tomando $\alpha = \frac{1}{40}$.

Figura 2.17: Regiones de accesibilidad de la raíz $z = 1$ de la ecuación $F(z) = z^3 - 1 = 0$ para el método simplificado de la secante (región roja) y el método tipo secante correspondiente a $\lambda = 0$ y tomando $\alpha = \frac{1}{64}$.

$$\begin{aligned}
\|I - L_0^{-1}A_1\| &\leq \|L_0^{-1}\|\,\|L_0 - A_1\| \\
&\leq \gamma K\left(\|\lambda z_1 + (1-\lambda)z_0 - z_{-1}\| + \|z_1 - z_0\|\right) \\
&\leq \gamma K\left((1+\lambda)\,\|z_1 - z_0\| + \|z_0 - z_{-1}\|\right) \\
&\leq \gamma K\left((1+\lambda)R + \alpha_0\right) \\
&< M,
\end{aligned}$$

$$\|A_1^{-1}\| = \|[\lambda z_1 + (1-\lambda)z_0, z_1; F]^{-1}\| \leq \frac{\gamma}{1-M} = \beta_1,$$

$$\begin{aligned}
\|F(z_1)\| &\leq \|L_1 - L_0\|\,\|z_1 - z_0\| \\
&\leq K\left(\|z_0 - z_{-1}\| + \|z_1 - z_0\|\right)\|z_1 - z_0\| \\
&\leq K(\alpha_0 + \varepsilon)\,\|z_1 - z_0\|,
\end{aligned}$$

$$\|z_2 - z_1\| \leq \|L_0^{-1}\|\,\|F(z_1)\| \leq \gamma K(\alpha_0 + \varepsilon)\eta_0 \leq \gamma K(2R + \alpha_0)\eta_0 = M\eta_0 = \eta_1,$$

siempre que $M < 1$. Además, $\eta_1 = M\eta_0 < \eta_0$ si $M < 1$.

Segundo paso del método predictor: definición de la terna $(\alpha_2, \beta_2, \eta_2)$.

Como

$$\|z_2 - z_1\| = \|L_0^{-1}F(z_1)\| \leq \|L_0^{-1}\|\,\|F(z_1)\| \leq \gamma K(\alpha_0 + \varepsilon)\eta_0 \leq \eta_1 = \alpha_2,$$

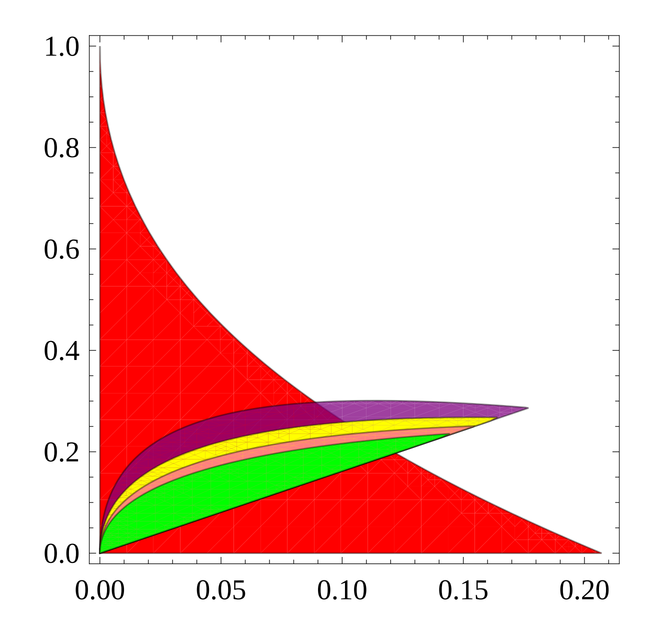


Figura 2.18: Dominios de parámetros del método simplificado de la secante (2.1) asociado al teorema 2.5 (región roja) y de los métodos tipo secante asociados al teorema 2.2 cuando $\lambda = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ (regiones verde, rosa, amarilla y morada, respectivamente).

$$\begin{aligned}
\|I - L_0^{-1}A_2\| &\leq \|L_0^{-1}\|\,\|L_0 - A_2\| \\
&\leq \gamma K\left(\|\lambda z_2 + (1-\lambda)z_1 - z_{-1}\| + \|z_2 - z_0\|\right) \\
&\leq \gamma K\left((1+\lambda)\|z_2 - z_0\| + (1-\lambda)\|z_1 - z_0\| + \|z_0 - z_{-1}\|\right) \\
&\leq \gamma K(2R + \alpha_0) \\
&= M,
\end{aligned}$$

$$\|A_2^{-1}\| = \left\|[\lambda z_2 + (1-\lambda)z_1, z_1; F]^{-1}\right\| \leq \frac{\gamma}{1-M} = \beta_2,$$

$$\begin{aligned}
\|F(z_2)\| &\leq \|L_2 - L_0\|\,\|z_2 - z_1\| \\
&\leq K\left(\|z_1 - z_{-1}\| + \|z_2 - z_0\|\right)\|z_2 - z_1\| \\
&\leq K(2R + \alpha_0)\,\|z_2 - z_1\|,
\end{aligned}$$

$$\|z_3 - z_2\| \leq \|L_0^{-1}\|\,\|F(z_2)\| \leq \gamma K(2R + \alpha_0)\|z_2 - z_1\| \leq M\eta_1 = \eta_2,$$

siempre que $M < 1$. Además, $\alpha_2 = \eta_1 = M\eta_0 = M\alpha_1 < \alpha_1$ y $\eta_2 = M\eta_1 = M^2\eta_0 < \eta_0$ si $M < 1$.

n-ésimo paso del método predictor: definición de la terna $(\alpha_n, \beta_n, \eta_n)$.

Como

$$\begin{aligned}
\|z_n - z_{n-1}\| &= \|L_0^{-1}F(z_{n-1})\| \leq \|L_0^{-1}\|\,\|F(z_{n-1})\| \\
&\leq \gamma K(2R + \alpha_0)\|z_{n-1} - z_{n-2}\| \\
&\leq \gamma K(2R + \alpha_0)\eta_{n-2} \\
&= M\eta_{n-2} = \eta_{n-1} = \alpha_n,
\end{aligned}$$

$$\begin{aligned} \|I - L_0^{-1}A_n\| &\leq \|L_0^{-1}\|\,\|L_0 - A_n\| \\ &\leq \gamma K\left(\|\lambda z_n + (1-\lambda)z_{n-1} - z_{-1}\| + \|z_n - z_0\|\right) \\ &\leq \gamma K\left((1+\lambda)\|z_n - z_0\| + (1-\lambda)\|z_{n-1} - z_0\| + \|z_0 - z_{-1}\|\right) \\ &\leq \gamma K(2R + \alpha_0) \\ &= M, \end{aligned}$$

$$\|A_n^{-1}\| = \left\|[\lambda z_n + (1-\lambda)z_{n-1}, z_n; F]^{-1}\right\| \leq \frac{\gamma}{1-M} = \beta_n,$$

$$\begin{aligned} \|F(z_n)\| &\leq \|L_n - L_0\|\,\|z_n - z_{n-1}\| \\ &\leq K\left(\|z_{n-1} - z_{-1}\| + \|z_n - z_0\|\right)\,\|z_n - z_{n-1}\| \\ &\leq K(2R + \alpha_0)\,\|z_n - z_{n-1}\|, \end{aligned}$$

$$\|z_{n+1} - z_n\| \leq \|L_0^{-1}\|\,\|F(z_n)\| \leq \gamma K(2R+\alpha_0)\|z_n - z_{n-1}\| \leq M\eta_{n-1} = \eta_n,$$

siempre que $M < 1$. Además, $\alpha_n = \eta_{n-1} = M^{n-1}\eta_0 = M^{n-1}\alpha_1 < \alpha_1$ y $\eta_n = M\eta_{n-1} = M^n\eta_0 < \eta_0$ si $M < 1$.

Una vez construida la terna $(\alpha_n, \beta_n, \eta_n)$, formada a partir de las sucesiones reales $\{\alpha_n\}$, $\{\beta_n\}$ y $\{\eta_n\}$, buscamos un valor $N_0 \in \mathbb{N}$, de manera que la terna $(\alpha_{N_0}, \beta_{N_0}, \eta_{N_0})$ verifique las condiciones de convergencia dadas en (2.5) para la familia de métodos tipo secante.

En primer lugar, consideramos la correspondiente primera condición de (2.5):

$$\frac{\eta_{N_0}}{\alpha_{N_0} + \eta_{N_0}} < \frac{3-\sqrt{5}}{2}.$$

Como $\eta_n = M\eta_{n-1}$ y $\alpha_n = \eta_{n-1}$, la condición anterior se transforma en

$$\frac{M}{1+M} < \frac{3-\sqrt{5}}{2} \qquad \Leftrightarrow \qquad M < \frac{\sqrt{5}-1}{2}.$$

Debido al tipo de sucesiones que hemos definido y las cotas que hemos utilizado, observamos que el cálculo de nuevas iteraciones con el método predictor no permite verificar la condición a partir de un n, luego esta condición tendremos que imponerla inicialmente. Notemos que esta condición no representa una restricción excesiva.

En segundo lugar, la correspondiente segunda condición de (2.5) se transforma en

$$K\beta_{N_0} < \frac{\eta_{N_0}}{(\alpha_{N_0} + \eta_{N_0})(\alpha_{N_0} + \eta_{N_0} - \lambda(\alpha_{N_0} - \eta_{N_0}))}.$$

Luego,

$$\frac{K\gamma}{1-M} < \frac{M\eta_{N_0-1}}{(1+M)\eta_{N_0-1}(1+M-\lambda(1-M))\eta_{N_0-1}} = \frac{M}{(1+M)(1+M-\lambda(1-M))\eta_{N_0-1}},$$

de manera que

$$K\gamma\eta_{N_0-1} < \frac{M(1-M)}{(1+M)(1+M-\lambda(1-M))}.$$

Así

$$M^{N_0-1}K\gamma\eta_0 < \frac{M(1-M)}{(1+M)(1+M-\lambda(1-M))}$$

y teniendo en cuenta $M<1$ y $\eta_0=\varepsilon$, escribimos

$$N_0 > 1 + \frac{\log\left(\frac{M(1-M)}{(1+M)(1+M-\lambda(1-M))}\right) - \log\left(K\gamma\varepsilon\right)}{\log M}.$$

En consecuencia, una vez fijado $\lambda\in[0,1)$ y denotado la parte entera del número real t por $[t]$, tomamos

$$N_0 = 1 + \left[1 + \frac{\log\left(\frac{M(1-M)}{(1+M)(1+M-\lambda(1-M))}\right) - \log\left(K\gamma\varepsilon\right)}{\log M}\right] \tag{2.15}$$

y ya podemos asegurar que los métodos tipo secante convergen cuando parten de los puntos $x_{-1}=z_{N_0-1}$ y $x_0=z_{N_0}$, donde N_0 está definido en (2.15).

Finalmente, una vez estimado a priori el valor de N_0, resumimos todo lo anterior en el siguiente resultado.

Teorema 2.6. *Sean X e Y dos espacios de Banach y $F:\Omega\subset X\to Y$ un operador definido en un conjunto abierto convexo no vacío Ω. Supongamos que se cumplen las condiciones (H1)–(H4) y (2.11). Si $B(z_0,R)\subset\Omega$, con R dado en (2.13), y $M=K\gamma(2R+\alpha_0)<\frac{\sqrt{5}-1}{2}$, entonces la sucesión dada por el método híbrido (2.14) está bien definida y converge a una solución de la ecuación $F(x)=0$, donde N_0 está definido en (2.15).*

2.4. Aplicación

Hemos justificado anteriormente que la aplicación de los métodos tipo secante es más restrictiva que la del método simplificado de la secante (2.1). Ilustrémoslo con un ejemplo. Consideramos la siguiente ecuación integral no lineal de tipo Hammerstein

$$x(s) = 1 + \frac{1}{2}\int_0^1 G(s,t)\,x(t)^2\,dt, \quad s\in[0,1], \tag{2.16}$$

donde $x\in C[0,1]$, $t\in[0,1]$ y G es la función de Green en $[0,1]\times[0,1]$.

A continuación, transformamos la ecuación integral (2.16) en un problema de dimensión finita, tal y como se hizo en la sección 1.6.1 del capítulo 1, de manera que obtenemos el siguiente sistema de ecuaciones no lineales:

$$F(\mathbf{x}) \equiv \mathbf{x} - \mathbf{1} - \frac{1}{2}A\,\hat{\mathbf{x}} = 0, \qquad F:\mathbb{R}^8 \longrightarrow \mathbb{R}^8, \tag{2.17}$$

donde

$$\mathbf{x}=(x_1,x_2,\ldots,x_8)^T, \qquad \mathbf{1}=(1,1,\ldots,1)^T, \qquad A=(a_{ij})_{i,j=1}^8, \qquad \hat{\mathbf{x}}=(x_1^2,x_2^2,\ldots,x_8^2)^T.$$

Además,

$$[\mathbf{u},\mathbf{v};F] = I - \frac{1}{2}A\,\mathrm{diag}\{\mathbf{z}\},$$

donde $\mathbf{z}=(z_1,z_2,\ldots,z_8)^T$ y $z_i=u_i+v_i$, para todo $i=1,2,\ldots,8$.

Eligiendo $\mathbf{z_{-1}} = (9/10, 9/10, \ldots, 9/10)^T$ y $\mathbf{z_0} = (1, 1, \ldots, 1)^T$ como puntos de salida y la norma del máximo, obtenemos $\alpha_0 = 0.1$, $\gamma = 1.1305\ldots$, $\varepsilon = 0.0687\ldots$, $K = 0.0617\ldots$, $K\gamma\alpha_0 = 0.0069\ldots$ y $K\gamma\varepsilon = 0.0047\ldots$ Por tanto, se puede ver que se verifican las dos condiciones de (2.11) y podemos aplicar entonces el método simplificado de la secante (2.1) para aproximar una solución del sistema (2.17). Por contra, no podemos utilizar los métodos tipo secante, ya que la primera condición de (2.5) no se satisface, puesto que

$$\frac{\eta}{\alpha+\eta} = 0.4072\ldots > \frac{3-\sqrt{5}}{2} = 0.3819\ldots,$$

donde $\alpha = \alpha_0$ y $\eta = \varepsilon$.

Por el teorema 2.5, el método simplificado de la secante (2.1) es convergente y, después de ocho iteraciones y usando el criterio de parada $\|\mathbf{z_n} - \mathbf{z_{n-1}}\| < 10^{-16}$, obtenemos la aproximación numérica $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_8^*)^T$ de la solución de (2.17) que vemos en la tabla 2.1. En la tabla 2.2 mostramos los errores $\|\mathbf{z_n} - \mathbf{x}^*\|$ obtenidos usando el mismo criterio de parada. Notemos que el vector dado en la tabla 2.1 es una buena aproximación de la solución del sistema (2.17), puesto que $\|F(\mathbf{x}^*)\| \leq \text{constante} \times 10^{-16}$. Mostramos la sucesión $\{\|F(\mathbf{z_n})\|\}$ en la tabla 2.2.

i	x_i^*	i	x_i^*	i	x_i^*	i	x_i^*
1	$1.005450\ldots$	3	$1.051629\ldots$	5	$1.069365\ldots$	7	$1.025815\ldots$
2	$1.025815\ldots$	4	$1.069365\ldots$	6	$1.051629\ldots$	8	$1.005450\ldots$

Tabla 2.1: Aproximación de la solución $\mathbf{x}^*$ de (2.17)

n	$\|\mathbf{z_n} - \mathbf{x}^*\|$	$\|F(\mathbf{z_n})\|$
-1	$1.6936\ldots \times 10^{-1}$	$1.5004\ldots \times 10^{-1}$
0	$6.9365\ldots \times 10^{-2}$	$6.1779\ldots \times 10^{-2}$
1	$6.6461\ldots \times 10^{-4}$	$5.9142\ldots \times 10^{-4}$
2	$8.6314\ldots \times 10^{-6}$	$7.6848\ldots \times 10^{-6}$
3	$1.1194\ldots \times 10^{-7}$	$9.9676\ldots \times 10^{-8}$
4	$1.4513\ldots \times 10^{-9}$	$1.2922\ldots \times 10^{-9}$
5	$1.8814\ldots \times 10^{-11}$	$1.6752\ldots \times 10^{-11}$
6	$2.4390\ldots \times 10^{-13}$	$2.1717\ldots \times 10^{-13}$
7	$3.1619\ldots \times 10^{-15}$	$2.8154\ldots \times 10^{-15}$

Tabla 2.2: Errores absolutos obtenidos con el método simplificado de la secante y $\{\|F(\mathbf{z_n})\|\}$

Además, por el teorema 2.5, la existencia de la solución está garantizada en $\overline{B(\mathbf{z_0}, 1.1683\ldots)}$ y es única en $B(\mathbf{z_0}, 13.0286\ldots)$.

Ahora, vamos a aplicar el método híbrido (2.14) con $\lambda = \frac{1}{2}$ para aproximar la solución de (2.17) dada en la tabla 2.1. Para esto, teniendo en cuenta que $M = K\gamma(2R + \alpha_0) = 0.0166\ldots < 0.6180\ldots = \frac{\sqrt{5}-1}{2}$, calculamos el valor N_0 determinado por el teorema 2.6. De acuerdo con la fórmula (2.15), $N_0 = 1$ para $\lambda = \frac{1}{2}$; por tanto, después de una iteración del

método simplificado de la secante (2.1), aplicamos el método tipo secante correspondiente a $\lambda = \frac{1}{2}$ y obtenemos la solución aproximada dada en la tabla 2.1 después de tres iteraciones más. En la tabla 2.3 se pueden ver los errores $\|\mathbf{x_n} - \mathbf{x}^*\|$ con el criterio de parada $\|\mathbf{x_n} - \mathbf{x_{n-1}}\| < 10^{-16}$, así como la sucesión $\{\|F(\mathbf{x_n})\|\}$.

n	$\|\mathbf{x_n} - \mathbf{x}^*\|$	$\|F(\mathbf{x_n})\|$
-1	$1.6936\ldots \times 10^{-1}$	$1.5004\ldots \times 10^{-1}$
0	$6.9365\ldots \times 10^{-2}$	$6.1779\ldots \times 10^{-2}$
1	$6.6461\ldots \times 10^{-4}$	$5.9142\ldots \times 10^{-4}$
2	$1.5148\ldots \times 10^{-10}$	$1.3499\ldots \times 10^{-10}$
3	$2.6106\ldots \times 10^{-15}$	$2.3266\ldots \times 10^{-15}$

Tabla 2.3: Errores absolutos obtenidos con el método híbrido (2.14) correspondiente a $\lambda = \frac{1}{2}$ y $\{\|F(\mathbf{x_n})\|\}$

Capítulo 3

Situación (no)-diferenciable

En este capítulo, vamos a estudiar la convergencia semilocal de los métodos tipo secante persiguiendo dos objetivos. En primer lugar, a partir de una modificación de la técnica de demostración dada por las relaciones de recurrencia en el teorema 2.2 del capítulo anterior, obtenemos un resultado de convergencia semilocal para la familia de métodos tipo secante que mejora notablemente el dominio de parámetros asociado al teorema 2.2. Además, el resultado así obtenido proporciona un resultado nuevo de convergencia semilocal cuando el operador implicado F es no diferenciable, situación habitualmente no contemplada y que se deduce gracias a la novedosa condición que se exige a la diferencia dividida de primer orden de F.

En la sección 3.1 recordamos el principal problema que presenta el estudio de la convergencia semilocal dado para la familia de métodos tipo secante en el capítulo anterior, que, tal y como hemos visto, es que limita mucho la aplicación de estos métodos para resolver $F(x) = 0$. Además, presentamos la novedosa condición que exigiremos a la diferencia dividida de primer orden de F. En la sección 3.2 vemos que el dominio de parámetros de la familia de métodos tipo secante se puede mejorar utilizando una modificación de la técnica de demostración de la convergencia semilocal presentada, que, además, como hemos dicho antes, va a permitir contemplar situaciones en las que el operador F es no diferenciable. Finalmente, en la sección 3.3, ilustramos lo anterior con dos sistemas no lineales, uno diferenciable y otro no diferenciable.

3.1. Planteamiento del problema

Hemos visto en el capítulo anterior, teorema 2.2, un resultado de convergencia semilocal bajo la exigencia de que exista la diferencia dividida de primer orden del operador F para cada par de puntos distintos en Ω y de que se cumplan las siguientes condiciones iniciales:

(C1) $\|x_0 - x_{-1}\| = \alpha \neq 0$ con $x_{-1}, x_0 \in \Omega$,

(C2) fijado $\lambda \in [0, 1)$, existe $A_0^{-1} = [y_0, x_0; F]^{-1} \in \mathcal{L}(Y, X)$, para $x_0, y_0 \in \Omega$, y es tal que $\|A_0^{-1}\| \leq \beta$,

(C3) $\|[y_0, x_0; F]^{-1} F(x_0)\| \leq \eta$,

(C4) $\|[x, y; F] - [u, v; F]\| \leq K(\|x - u\| + \|y - v\|)$, $\quad K \geq 0$, $x, y, u, v \in \Omega$, $\quad x \neq y$, $u \neq v$.

Se obtienen entonces condiciones suficientes para que se dé la convergencia semilocal de la familia de métodos tipo secante, tal y como hemos visto en el teorema 2.2 del capítulo

anterior. Este resultado requiere que el operador F sea diferenciable, además de tener asociado un dominio de parámetros reducido y, junto con los inconvenientes descritos en el capítulo anterior, implica una deficiente aplicabilidad de los métodos tipo secante.

Si consideramos ecuaciones integrales no lineales de tipo Hammerstein de la forma (1.41), presentadas en la sección 1.6.1, donde

$$H(t, x(t)) = \delta x(t)^2 + \mu |x(t)|, \quad \delta, \mu \in \mathbb{R},$$

y las trasformamos, mediante un proceso de discretización, en sistemas de ecuaciones no lineales de la forma

$$\Psi(\mathbf{x}) \equiv \mathbf{x} - \mathbf{f} - A\,(\delta\,\hat{\mathbf{x}} + \mu\,\tilde{\mathbf{x}}) = 0, \qquad \Psi : \mathbb{R}^m \longrightarrow \mathbb{R}^m, \tag{3.1}$$

donde $\hat{\mathbf{x}} = (x_1^2, x_2^2, \ldots, x_m^2)^T$, $\tilde{\mathbf{x}} = (|x_1|, |x_2|, \ldots, |x_m|)^T$ y $\delta, \mu \in \mathbb{R}$, es obvio que la función Ψ definida en (3.1) es no lineal y además no diferenciable si $\mu \neq 0$. Teniendo en cuenta (1.46), consideramos

$$[\mathbf{u}, \mathbf{v}; \Psi] = I - A\,(\delta \mathrm{diag}\{\mathbf{z}\} + \mu \mathrm{diag}\{\mathbf{w}\}),$$

donde $\mathbf{z} = (z_1, z_2, \ldots, z_m)^T$ con $z_i = u_i + v_i$, para todo $i = 1, 2, \ldots, m$, y $\mathbf{w} = (w_1, w_2, \ldots, w_m)^T$ con $w_i = \frac{|u_i| - |v_i|}{u_i - v_i}$, para todo $i = 1, 2, \ldots, m$, de manera que

$$\|[\mathbf{x}, \mathbf{y}; \Psi] - [\mathbf{u}, \mathbf{v}; \Psi]\| \leq L + K(\|\mathbf{x} - \mathbf{u}\| + \|\mathbf{y} - \mathbf{v}\|) \text{ con } L = 2|\mu|\,\|A\| \text{ y } K = |\delta|\,\|A\|. \tag{3.2}$$

Por tanto, si en vez de exigir la condición (C4) para la diferencia dividida de primer orden del operador F, exigimos que F cumpla una condición del tipo (3.2) en el espacio de Banach X, la dada por

$$\|[x, y; F] - [u, v; F]\| \leq L + K(\|x - u\| + \|y - v\|); L, K \geq 0; x, y, u, v \in \Omega; x \neq y; u \neq v, \tag{3.3}$$

podremos considerar situaciones en las que el operador F sea no diferenciable ($L \neq 0$), como por ejemplo la indicada anteriormente para la función Ψ definida en (3.1) con $\mu \neq 0$. Por el contrario, para $L = 0$, podemos seguir considerando situaciones en las que el operador F sea diferenciable, lo que se contempla en (3.3) si $\mu = 0$.

Así, para el caso anterior, vamos a obtener un nuevo resultado de convergencia semilocal exigiendo la condición (3.3) al operador F, resultado que para $L = 0$ permitirá mejorar los dominios de parámetros asociados al teorema 2.2 para los métodos tipo secante cuando el operador F sea diferenciable, mientras que para $L \neq 0$ permitirá obtener un resultado de convergencia semilocal en situaciones en las que el operador F sea no diferenciable.

3.2. Mejora de la accesibilidad

Nuestro primer objetivo es entonces ampliar el dominio de parámetros de los métodos tipo secante. Para ello desarrollamos una nueva técnica para analizar la convergencia semilocal de estos métodos, que además, como segundo objetivo, va a permitir utilizar los métodos tipo secante para resolver ecuaciones en situaciones en las que el operador F sea no diferenciable. Como en el teorema 2.2, exigiremos que exista una diferencia dividida de primer orden del operador F para cada par de puntos distintos de Ω.

Por otra parte, si la sucesión $\{x_n\}$ está generada por los métodos tipo secante, resulta evidente en el estudio de su convergencia que existirán todas las diferencias divididas de

primer orden $[y_k, x_k; F]$, salvo que $x_k = y_k = \lambda x_k + (1-\lambda)x_{k-1}$, en cuyo caso resulta evidente que $x_k = x_{k-1}$ y, por tanto, $x_{k-1} = x_k = x^*$ es una solución de la ecuación $F(x) = 0$ y $x_n = x^*$, para todo $n \geq k-1$, de manera que la sucesión $\{x_n\}$ es convergente a la solución x^* de $F(x) = 0$.

Teniendo en cuenta estas ideas, comenzamos dando un lema técnico que utilizamos posteriormente.

Lema 3.1. *Sea $\{x_n\}$ la sucesión dada por los métodos tipo secante. Si $x_{m-1} \neq x_m$ con $x_{m-1}, x_m \in \Omega$, entonces*

$$F(x_m) = ([x_m, x_{m-1}; F] - A_{m-1})(x_m - x_{m-1}), \quad \textit{donde} \quad A_{m-1} = [y_{m-1}, x_{m-1}; F].$$

DEMOSTRACIÓN. A partir de la definición de la sucesión $\{x_n\}$ se sigue

$$F(x_m) + [y_{m-1}, x_{m-1}; F](x_m - x_{m-1}) = 0,$$

de manera que

$$\begin{aligned} F(x_m) &= F(x_m) - F(x_{m-1}) - [y_{m-1}, x_{m-1}; F](x_m - x_{m-1}) \\ &= [x_m, x_{m-1}; F](x_m - x_{m-1}) - [y_{m-1}, x_{m-1}; F](x_m - x_{m-1}) \\ &= ([x_m, x_{m-1}; F] - A_{m-1})(x_m - x_{m-1}). \quad \blacksquare \end{aligned}$$

A continuación, presentamos el nuevo resultado de convergencia semilocal para los métodos tipo secante.

Teorema 3.2. *Sean X e Y dos espacios de Banach y $F : \Omega \subset X \to Y$ un operador no lineal definido en un dominio abierto convexo no vacío Ω. Supongamos que se cumplen las condiciones (C1)–(C3) y (3.3). Fijado $\lambda \in [0,1)$, si la ecuación*

$$h(t) = t\left(1 - \frac{m}{1 - \beta\,(L + K(2t + (1-\lambda)\alpha))}\right) - \eta = 0, \tag{3.4}$$

donde $m = \text{máx}\,\{\,\beta(L + K((1-\lambda)\alpha + \eta)),\ \beta(L + (2-\lambda)K\eta)\,\}$, tiene al menos una raíz positiva, y denotamos por R la raíz positiva más pequeña de (3.4),

$$\beta(L + K(2R + (1-\lambda)\alpha)) < 1 \tag{3.5}$$

y $B(x_0, R) \subset \Omega$, entonces la sucesión $\{x_n\}$ dada por los métodos tipo secante, empezando en x_{-1} y x_0, está bien definida y converge a una solución x^* de $F(x) = 0$. Además, la solución x^* y las iteraciones x_n pertenecen a $\overline{B(x_0, R)}$ y x^* es única en $\overline{B(x_0, R)} \cap \Omega$.

DEMOSTRACIÓN. Comenzamos probando que la sucesión $\{x_n\}$ está bien definida, es decir, $x_n \in B(x_0, R) \subset \Omega$, para todo $n \in \mathbb{N}$. A partir de (C3) se sigue que x_1 está bien definido y $\|x_1 - x_0\| = \|A_0^{-1}F(x_0)\| \leq \eta < R$, por ser R solución de (3.4). Luego, $x_1 \in B(x_0, R)$.

A continuación, utilizando (3.3), vemos que

$$\begin{aligned} \|I - A_0^{-1}A_1\| &\leq \|A_0^{-1}\|\|A_0 - A_1\| \\ &\leq \beta(L + K(\|y_1 - y_0\| + \|x_1 - x_0\|)) \\ &\leq \beta(L + K((1+\lambda)\|x_1 - x_0\| + (1-\lambda)\|x_0 - x_{-1}\|)) \\ &\leq \beta(L + K((1+\lambda)R + (1-\lambda)\alpha)) \\ &\leq \beta(L + K(2R + (1-\lambda)\alpha)) \\ &< 1, \end{aligned}$$

y por el lema de Banach (lema 1.22), tenemos que existe A_1^{-1} y es tal que

$$\|A_1^{-1}\| \leq \frac{\beta}{1-\beta(L+K(2R+(1-\lambda)\alpha))}.$$

Además, por el lema 3.1, se sigue $F(x_1) = ([x_1, x_0; F] - A_0)(x_1 - x_0)$, de manera que

$$\begin{aligned}
\|F(x_1)\| &\leq \|[x_1, x_0; F] - A_0\| \, \|x_1 - x_0\| \\
&\leq (L + K\|y_0 - x_1\|)\|x_1 - x_0\| \\
&\leq (L + K(\|y_0 - x_0\| + \|x_1 - x_0\|))\|x_1 - x_0\| \\
&\leq (L + K((1-\lambda)\alpha + \eta))\|x_1 - x_0\|.
\end{aligned}$$

En consecuencia,

$$\begin{aligned}
\|x_2 - x_1\| &\leq \|A_1^{-1}\| \, \|F(x_1)\| \\
&\leq \frac{\beta(L + K((1-\lambda)\alpha + \eta))}{1-\beta(L+K(2R+(1-\lambda)\alpha))} \|x_1 - x_0\| \\
&\leq P\|x_1 - x_0\| \\
&< \eta,
\end{aligned}$$

donde $P = \dfrac{m}{1-\beta(L+K(2R+(1-\lambda)\alpha))} < 1$, y

$$\|x_2 - x_0\| \leq \|x_2 - x_1\| + \|x_1 - x_0\| \leq (1+P)\|x_1 - x_0\| = \frac{1-P^2}{1-P}\|x_1 - x_0\| < \frac{\eta}{1-P} = R.$$

Por lo tanto, $x_2 \in B(x_0, R)$.

Ahora, podemos demostrar por inducción matemática sobre n que se cumplen los siguientes tres ítems para $n \in \mathbb{N}$:

- El operador A_{n-1}^{-1} existe y es tal que $\|A_{n-1}^{-1}\| \leq \dfrac{\beta}{1-\beta(L+K(2R+(1-\lambda)\alpha))}$,
- $\|x_n - x_{n-1}\| \leq P\,\|x_{n-1} - x_{n-2}\| \leq P^{n-1}\,\|x_1 - x_0\| < \eta$,
- $\|x_{n+1} - x_0\| \leq \dfrac{1-P^{n+1}}{1-P}\|x_1 - x_0\| < \dfrac{\eta}{1-P} = R$,

siempre que $A_i = [y_i, x_i; F]$ sea inversible y $x_{i+1} \in B(x_0, R)$, $\forall i = 1, 2, \ldots, n-2$.

En primer lugar, por hipótesis, vemos que

$$\begin{aligned}
\|I - A_0^{-1}A_n\| &\leq \|A_0^{-1}\|\|A_0 - A_n\| \\
&\leq \beta(L + K(\|y_n - y_0\| + \|x_n - x_0\|)) \\
&\leq \beta(L + K(\lambda\|x_n - x_0\| + (1-\lambda)\|x_{n-1} - x_{-1}\| + \|x_n - x_0\|)) \\
&\leq \beta(L + K((1+\lambda)\|x_n - x_0\| + (1-\lambda)\|x_{n-1} - x_{-1}\|)) \\
&\leq \beta(L + K((1+\lambda)\|x_n - x_0\| + (1-\lambda)\|x_{n-1} - x_0\| + (1-\lambda)\|x_0 - x_{-1}\|)) \\
&< \beta(L + K(2R + (1-\lambda)\alpha)) \\
&< 1.
\end{aligned}$$

Además,

$$\|A_n^{-1}\| \leq \frac{\beta}{1-\beta(L+K(2R+(1-\lambda)\alpha))}.$$

En segundo lugar, por el lema 3.1, tenemos $F(x_n) = ([x_n, x_{n-1}; F] - A_{n-1})(x_n - x_{n-1})$, de manera que

$$\begin{aligned}
\|F(x_n)\| &\leq \|[x_n, x_{n-1}; F] - A_{n-1}\| \, \|x_n - x_{n-1}\| \\
&\leq (L + K\|y_{n-1} - x_{n-1}\|) \, \|x_n - x_{n-1}\| \\
&\leq (L + K(\|y_{n-1} - x_{n-1}\| + \|x_n - x_{n-1}\|)) \, \|x_n - x_{n-1}\| \\
&\leq (L + K((1-\lambda)\|x_{n-1} - x_{n-2}\| + \|x_n - x_{n-1}\|)) \, \|x_n - x_{n-1}\| \\
&\leq (L + (2-\lambda)K\eta) \, \|x_n - x_{n-1}\|.
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
\|x_{n+1} - x_n\| &\leq \|A_n^{-1}\| \, \|F(x_n)\| \\
&\leq \frac{\beta(L+(2-\lambda)K\eta)}{1-\beta(L+K(2R+(1-\lambda)\alpha))} \|x_n - x_{n-1}\| \\
&\leq \frac{m}{1-\beta(L+K(2R+(1-\lambda)\alpha))} \|x_n - x_{n-1}\| \\
&= P \, \|x_n - x_{n-1}\| \\
&\leq P^n \, \|x_1 - x_0\| \\
&< \eta,
\end{aligned}$$

y, en consecuencia,

$$\begin{aligned}
\|x_{n+1} - x_0\| &\leq \sum_{i=0}^{n+1} \|x_i - x_{i-1}\| \leq (P^n + P^{n-1} + \cdots + P + 1)\|x_1 - x_0\| \\
&\leq \frac{1-P^{n+1}}{1-P} \|x_1 - x_0\| \\
&< \frac{\eta}{1-P} \\
&= R,
\end{aligned}$$

por ser $P < 1$. Luego $x_{n+1} \in B(x_0, R)$, para todo $n \in \mathbb{N}$.

Una vez probado que la sucesión $\{x_n\}$ está bien definida, vemos que es una sucesión de Cauchy. En efecto, como

$$\begin{aligned}
\|x_{n+j} - x_n\| &\leq \|x_{n+j} - x_{n+j-1}\| + \|x_{n+j-1} - x_{n+j-2}\| + \cdots + \|x_{n+1} - x_n\| \\
&\leq (P^{j-1} + P^{j-2} + \cdots + P + 1)\|x_{n+1} - x_n\| \\
&= \frac{1-P^j}{1-P} \|x_{n+1} - x_n\| \\
&< \frac{P^n}{1-P} \|x_1 - x_0\|,
\end{aligned}$$

para $j \geq 1$, y $P < 1$, es claro que $\{x_n\}$ es una sucesión de Cauchy. Por tanto, la sucesión $\{x_n\}$ es convergente.

Ahora, si $\lim_{n\to\infty} x_n = x^*$, vemos que x^* es una solución de $F(x) = 0$. Como

$$\|F(x_n)\| \leq (L + (2-\lambda)K\eta)\|x_n - x_{n-1}\|$$

y $\|x_n - x_{n-1}\| \to 0$ cuando $n \to \infty$, por la continuidad del operador F, se sigue fácilmente que $F(x^*) = 0$.

Finalmente, probamos la unicidad de la solución x^* en $\overline{B(x_0, R)}$. Suponemos que tenemos otra solución distinta $y^* \in \overline{B(x_0, R)}$ de la ecuación $F(x) = 0$. Sea $J = [y^*, x^*; F]$. Si J es inversible, tenemos que $x^* = y^*$, puesto que $J(y^* - x^*) = F(y^*) - F(x^*)$. Para ver que J es inversible, por el lema de Banach (lema 1.22), basta con ver que $\|I - A_0^{-1}J\| < 1$. En efecto,

$$\begin{aligned}
\|I - A_0^{-1}J\| &\leq \|A_0^{-1}\|\,\|A_0 - J\| \\
&\leq \|A_0^{-1}\|\,\|[y_0, x_0; F] - [y^*, x^*; F]\| \\
&\leq \beta(L + K(\|y^* - y_0\| + \|x^* - x_0\|)) \\
&\leq \beta(L + K(\|y^* - x_0\| + \|x_0 - y_0\| + \|x^* - x_0\|)) \\
&\leq \beta(L + K(R + (1-\lambda)\alpha + R)) \\
&= \beta(L + K(2R + (1-\lambda)\alpha)) \\
&< 1.
\end{aligned}$$

Luego, el operador J^{-1} existe. ■

Notemos que existen las diferencias divididas utilizadas $[x_j, x_{j-1}; F]$ y $[y_{j-1}, x_{j-1}; F]$, ya que si existe $j \in \mathbb{N}$ tal que $x_j = x_{j-1}$, entonces $x_{j-1} = x^*$ (la solución buscada). Además, si $x_{j-1} = y_{j-1} = \lambda x_{j-1} + (1-\lambda)x_{j-2}$, entonces $x_{j-1} = x_{j-2}$, de manera que $x_{j-2} = x^*$. En ambos casos, la sucesión $\{x_n\}$ es convergente trivialmente.

Una vez probada la convergencia semilocal de la familia de métodos tipo secante, nuestro siguiente objetivo es ver cuál es el dominio de parámetros de esta familia. Para ello, trasformamos la ecuación (3.4) en la siguiente ecuación cuadrática equivalente:

$$2K\beta t^2 + (\beta(L + K((1-\lambda)\alpha - 2\eta) + m - 1)\, t + \eta\,(1 - \beta(L + (1-\lambda)K\alpha)) = 0, \qquad (3.6)$$

y vemos cuándo tiene raíces reales positivas. La ecuación anterior tendrá dos raíces reales positivas si

$$\beta(L + K((1-\lambda)\alpha - 2\eta) + m - 1 < 0, \qquad (3.7)$$

$$1 - \beta(L + (1-\lambda)K\alpha) > 0, \qquad (3.8)$$

$$\begin{aligned}
\Delta &= (\beta(L + K((1-\lambda)\alpha - 2\eta)) + m - 1)^2 - 8K\beta\eta\,(1 - \beta(L + (1-\lambda)K\alpha)) \\
&= \Big(\beta(L + K((1-\lambda)\alpha - 2\eta)) + m - 1 + \sqrt{8K\beta\eta\,(1 - \beta(L + (1-\lambda)K\alpha))}\Big) \\
&\quad \times \Big(\beta(L + K((1-\lambda)\alpha - 2\eta)) + m - 1 - \sqrt{8K\beta\eta\,(1 - \beta(L + (1-\lambda)K\alpha))}\Big) > 0.
\end{aligned}$$

Observamos que los dos factores de Δ son < 0 si

$$\beta(L + K((1-\lambda)\alpha - 2\eta)) + m + \sqrt{8K\beta\eta\,(1 - \beta(L + (1-\lambda)K\alpha))} < 1. \qquad (3.9)$$

Además, se cumple (3.7) si se satisface (3.9). Por lo tanto, la ecuación (3.6) tendrá dos raíces reales positivas si se cumplen (3.8) y (3.9). Si la ecuación (3.6) tiene dos raíces reales positivas,

entonces la raíz positiva más pequeña es:

$$\begin{aligned} R &= \frac{1}{4K\beta}\left(1-m-\beta(L+K((1-\lambda)\alpha-2\eta))\right. \\ &\left.-\sqrt{(\beta(L+K((1-\lambda)\alpha-2\eta)+m-1)^2-8K\beta\eta\left(1-\beta(L+(1-\lambda)K\alpha)\right)}\right). \end{aligned} \tag{3.10}$$

Notemos que también podemos considerar que la ecuación (3.6) tenga una raíz doble sin más que tener en cuenta la desigualdad no estricta en (3.9).

A continuación, vemos cuándo se cumple la condición (3.5) que aparece en el teorema 3.2. Para ello, distinguimos dos casos:

- $\alpha \geq \eta \Rightarrow m = \beta(L+K((1-\lambda)\alpha+\eta))$,
- $\alpha \leq \eta \Rightarrow m = \beta(L+(2-\lambda)K\eta)$.

Si $\alpha \geq \eta$, entonces

$$\begin{aligned} R &= \frac{1}{4K\beta}\left(1-\beta(2L+K(2(1-\lambda)\alpha-\eta))\right. \\ &\left.-\sqrt{(\beta(2L+K(2(1-\lambda)\alpha-\eta))-1)^2-8K\beta\eta\left(1-\beta(L+(1-\lambda)K\alpha)\right)}\right) \end{aligned}$$

y, de (3.5), se sigue

$$\frac{1}{2}\left(1+K\beta\eta-\sqrt{(\beta(2L+K(2(1-\lambda)\alpha-\eta))-1)^2-8K\beta\eta\left(1-\beta(L+(1-\lambda)K\alpha)\right)}\right)<1,$$

$$\sqrt{(\beta(2L+K(2(1-\lambda)\alpha-\eta))-1)^2-8K\beta\eta\left(1-\beta(L+(1-\lambda)K\alpha)\right)}>K\beta\eta-1.$$

Para que se cumpla la última desigualdad, basta con que se cumpla $K\beta\eta-1<0$; es decir,

$$K\beta\eta<1. \tag{3.11}$$

Por lo tanto, la condición (3.5) del teorema 3.2 se cumple si se verifica (3.11) cuando $\alpha \geq \eta$.

Por otra parte, si $\alpha \leq \eta$, entonces

$$\begin{aligned} R &= \frac{1}{4K\beta}\left(1-\beta(2L+K((1-\lambda)\alpha-\lambda\eta))\right. \\ &\left.-\sqrt{(\beta(2L+K((1-\lambda)\alpha-\lambda\eta))-1)^2-8K\beta\eta\left(1-\beta(L+(1-\lambda)K\alpha)\right)}\right) \end{aligned}$$

y, de (3.5), se sigue

$$\begin{aligned} &\frac{1}{2}\left(1+K\beta((1-\lambda)\alpha+\eta)\right. \\ &\left.-\sqrt{(\beta(2L+K((1-\lambda)\alpha-\eta))-1)^2-8K\beta\eta\left(1-\beta(L+(1-\lambda)K\alpha)\right)}\right)<1, \end{aligned}$$

$$\sqrt{(\beta(2L+K((1-\lambda)\alpha-\eta))-1)^2-8K\beta\eta\left(1-\beta(L+(1-\lambda)K\alpha)\right)}>K\beta((1-\lambda)\alpha+\eta)-1.$$

Para que se cumpla la desigualdad anterior basta con que se cumpla $K\beta((1-\lambda)\alpha+\eta)-1<0$; es decir,

$$K\beta((1-\lambda)\alpha+\eta)<1. \tag{3.12}$$

Por lo tanto, la condición (3.5) del teorema 3.2 se cumple si se verifica (3.12) cuando $\alpha \leq \eta$.

Destacamos que no consideramos la posibilidad de que la ecuación cuadrática (3.4) tenga una raíz real positiva y otra negativa porque, en este caso, la raíz real positiva nunca cumple la condición (3.5).

En consecuencia, las condiciones de convergencia que se imponen a los parámetros α, β, η, L y K, como son que la ecuación (3.4) tenga al menos una raíz real positiva y que la raíz positiva más pequeña de (3.4), denotada por R, cumpla (3.5), se van a cumplir siempre que se verifiquen (3.8), (3.9) y (3.11) cuando $\alpha \geq \eta$ o (3.8), (3.9) y (3.12) cuando $\alpha \leq \eta$. Así, enunciamos entonces el siguiente resultado, cuya demostración se sigue fácilmente sin más que satisfacer las hipótesis del teorema 3.2.

Corolario 3.3. *Sean X e Y dos espacios de Banach y $F : \Omega \subset X \to Y$ un operador no lineal definido en un dominio abierto convexo no vacío Ω. Una vez fijado $\lambda \in [0, 1)$, suponemos que se cumplen las condiciones (C1)–(C3), (3.3), (3.8) y (3.9). Además, si se cumplen (3.11) cuando $\alpha \geq \eta$ o (3.12) cuando $\alpha \leq \eta$ y $B(x_0, R) \subset \Omega$, donde R está definido en (3.10) con $m = \max\{\beta(L+K((1-\lambda)\alpha+\eta)), \beta(L+(2-\lambda)K\eta)\}$, entonces la sucesión $\{x_n\}$ dada por los métodos tipo secante, empezando en x_{-1} y x_0, está bien definida y converge a una solución x^* de $F(x) = 0$. Además, la solución x^* y las iteraciones x_n pertenecen a $\overline{B(x_0, R)}$ y x^* es única en $\overline{B(x_0, R)} \cap \Omega$.*

Ahora, representamos gráficamente el dominio de parámetros asociado al corolario anterior. Para representarlo gráficamente, se colorean en un plano xy, con $x = K\beta\eta$ e $y = K\beta\alpha$, los valores de los parámetros que verifican las condiciones (3.8), (3.9) y (3.11), si $\alpha \geq \eta$, o (3.12), si $\alpha \leq \eta$, que se imponen en el corolario anterior. Observamos que las condiciones exigidas a los puntos de salida, (C1)–(C3), introducen los parámetros α, β y η, y la condición exigida al operador F, (3.3), introduce los parámetros fijos L y K. A continuación, distinguimos dos casos: el caso diferenciable, $L = 0$, y el no diferenciable, $L \neq 0$.

Comenzamos con el caso diferenciable. Como se puede ver en las figuras 3.1 y 3.2, el nuevo dominio de parámetros (figura 3.1) resuelve el problema que tienen los métodos tipo secante en el dominio de parámetros asociado al teorema 2.2 (figura 3.3) cuando los valores de η son pequeños para un valor fijo de α. Fijado un valor de $K\beta\alpha$ perteneciente al dominio de parámetros, observamos que el valor de $K\beta\eta$ sólo está acotado superiormente y no inferiormente. Notamos que las regiones están superpuestas.

En segundo lugar, representamos el dominio de parámetros asociado al teorema 3.2 cuando F es no diferenciable, $L \neq 0$. En este caso, el dominio de parámetros depende de la cantidad $L\beta$, que tiene que cumplir que $L\beta < \frac{1}{2}$, sin más que observar la condición (3.9). Así, representamos gráficamente los casos $L\beta = \frac{1}{10}$ y $L\beta = \frac{1}{5}$ en las figuras 3.4 y 3.5, respectivamente. Observamos que el dominio de parámetros se reduce al aumentar el valor de $L\beta$, pero sigue teniendo buenas características locales. Notamos que las regiones están superpuestas.

3.3. Aplicación

A continuación ilustramos todo lo visto en la sección anterior con dos sistemas de ecuaciones no lineales. En primer lugar, consideramos un sistema de ecuaciones no lineales diferenciable y vemos que no podemos garantizar la convergencia semilocal de los métodos tipo secante mediante el teorema 2.2, pero si que lo podemos hacer mediante el teorema 3.2. En

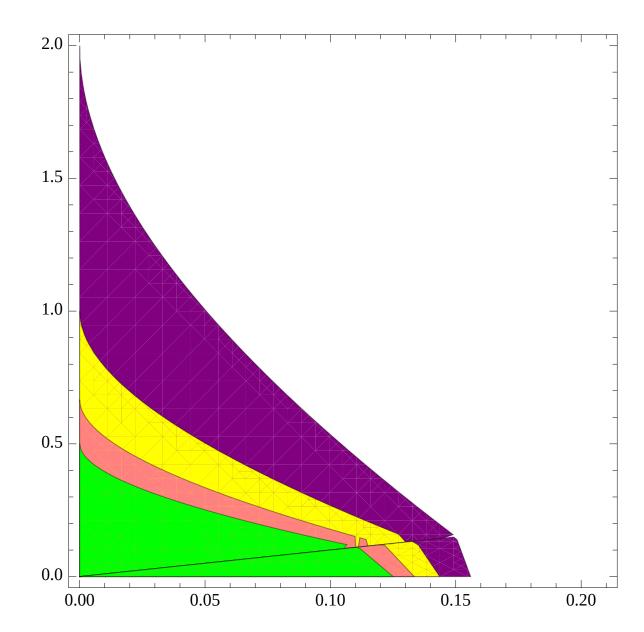


Figura 3.1: Dominios de parámetros de los métodos tipo secante asociados al corolario 3.3 cuando $L = 0$ y $\lambda = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ (verde, rosa, amarillo y morado, respectivamente).

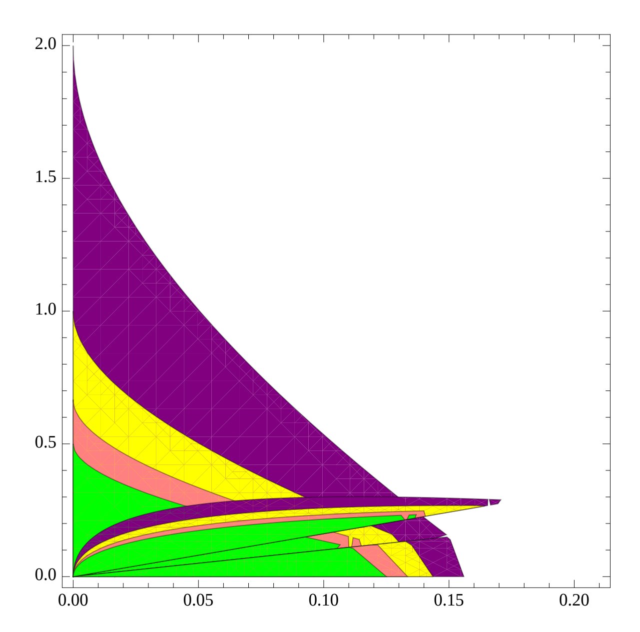


Figura 3.2: Dominios de parámetros de los métodos tipo secante asociados al teorema 2.2 y al corolario 3.3 cuando $L = 0$ y $\lambda = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ (verde, rosa, amarillo y morado, respectivamente).

segundo lugar, consideramos un sistema de ecuaciones no lineales no diferenciable y vemos que el teorema 3.2 garantiza la convergencia semilocal de los métodos tipo secante.

3.3.1. Sistema de ecuaciones no lineales diferenciable

Sea la siguiente ecuación integral no lineal de tipo Hammerstein

$$x(s) = 1 + \int_0^1 G(s,t)\, x(t)^2\, dt, \quad s \in [0,1], \tag{3.13}$$

donde $x \in C[0,1]$ y G es la función de Green en $[0,1] \times [0,1]$.

A continuación, mediante un proceso de discretización, transformamos (3.13) en un problema de dimensión finita, tal y como se hizo en la sección 1.6.1 del capítulo 1, de manera que obtenemos el siguiente sistema de ecuaciones no lineales:

$$F(\mathbf{x}) \equiv \mathbf{x} - \mathbf{1} - A\hat{\mathbf{x}} = 0, \qquad F : \mathbb{R}^8 \longrightarrow \mathbb{R}^8, \tag{3.14}$$

donde

$$\mathbf{x} = (x_1, x_2, \dots, x_8)^T, \qquad \mathbf{1} = (1, 1, \dots, 1)^T, \qquad A = (a_{ij})_{i,j=1}^8, \qquad \hat{\mathbf{x}} = (x_1^2, x_2^2, \dots, x_8^2)^T.$$

En este caso, el operador diferencia dividida considerado es

$$[\mathbf{u}, \mathbf{v}; F] = I - \frac{1}{2} A\, \mathrm{diag}\{\mathbf{z}\},$$

donde $\mathbf{z} = (z_1, z_2, \dots, z_8)^T$ y $z_i = u_i + v_i$, para todo $i = 1, 2, \dots, 8$.

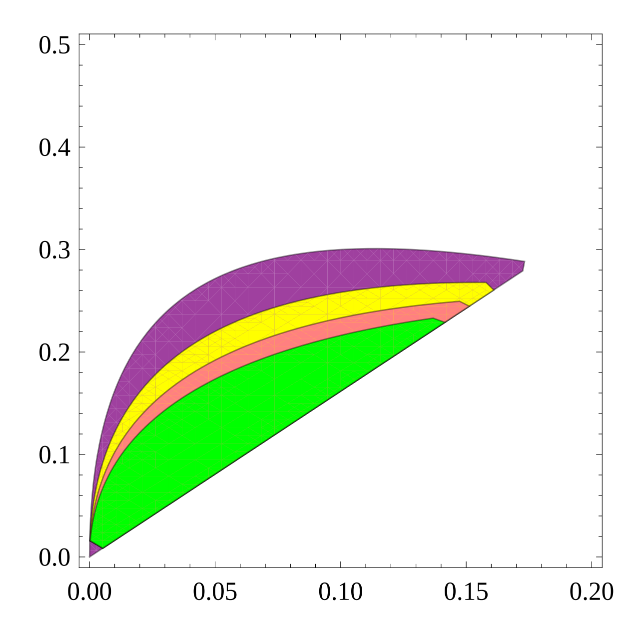


Figura 3.3: Dominios de parámetros de los métodos tipo secante asociados al teorema 2.2 cuando $\lambda = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ (regiones verde, rosa, amarilla y morada, respectivamente).

Eligiendo $\mathbf{x_{-1}} = \left(\frac{9}{10}, \frac{9}{10}, \dots, \frac{9}{10}\right)^T$ y $\mathbf{x_0} = (1, 1, \dots, 1)^T$ como puntos de salida, $\lambda = \frac{1}{2}$ y la norma del máximo, obtenemos $\alpha = 0.1$, $\beta = 1.3035\dots$, $\eta = 0.1556\dots$ y $K = 0.1235\dots$ Vemos entonces que la primera condición de (2.5) del teorema 2.2 no se cumple puesto que

$$\frac{\eta}{\alpha + \eta} = 0.6088\dots > \frac{3 - \sqrt{5}}{2} = 0.3819\dots$$

En consecuencia, no podemos garantizar la convergencia del método tipo secante con $\lambda = \frac{1}{2}$ mediante el teorema 2.2 del capítulo anterior. Sin embargo, si que la podemos garantizar mediante el teorema 3.2, puesto que la correspondiente ecuación (3.4) es

$$t\left(1 - \frac{(0.0376\dots)}{1 - (0.1610\dots)(2t + 0.05)}\right) - (0.1556\dots) = 0,$$

que tiene dos raíces reales positivas y la más pequeña, denotada por $R = 0.1621\dots$, cumple la condición (3.5):

$$\beta K(2R + (1 - \lambda)\alpha) = 0.0602\dots < 1.$$

En este caso, utilizando el método tipo secante correspondiente a $\lambda = \frac{1}{2}$, después de cinco iteraciones y usando el criterio de parada $\|\mathbf{x_n} - \mathbf{x_{n-1}}\| < 10^{-16}$, obtenemos la aproximación numérica $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_8^*)^T$ de una solución de (3.14) que vemos en la tabla 3.1. En la tabla 3.2 mostramos los errores $\|\mathbf{x_n} - \mathbf{x}^*\|$ obtenidos usando el mismo criterio de parada. Notemos que el vector dado en la tabla 3.1 es una buena aproximación de la solución del sistema (3.14), puesto que $\|F(\mathbf{x}^*)\| \leq \text{constante} \times 10^{-16}$. Mostramos la sucesión $\{\|F(\mathbf{x_n})\|\}$ en la tabla 3.2.

Además, por el teorema 3.2, la existencia y unicidad de la solución está garantizada en la bola $\overline{B(\mathbf{x_0}, 0.1621\dots)}$.

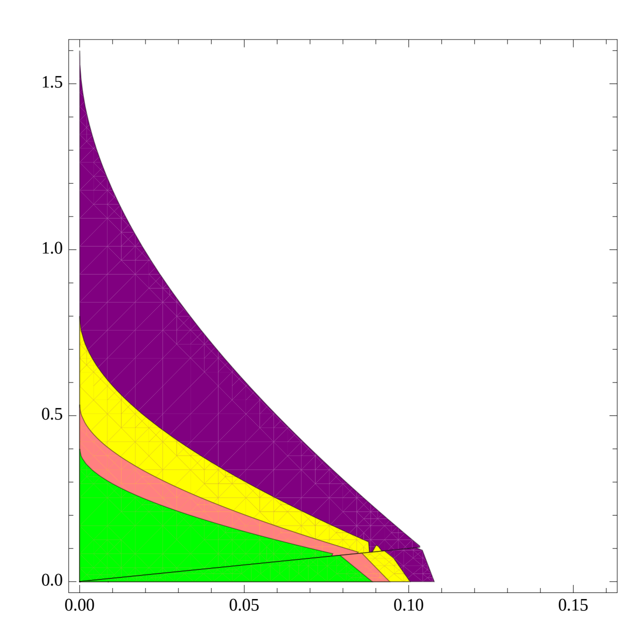


Figura 3.4: Dominios de parámetros de los métodos tipo secante asociados al corolario 3.3 cuando $L\beta = \frac{1}{10}$ y $\lambda = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ (verde, rosa, amarillo y morado, respectivamente).

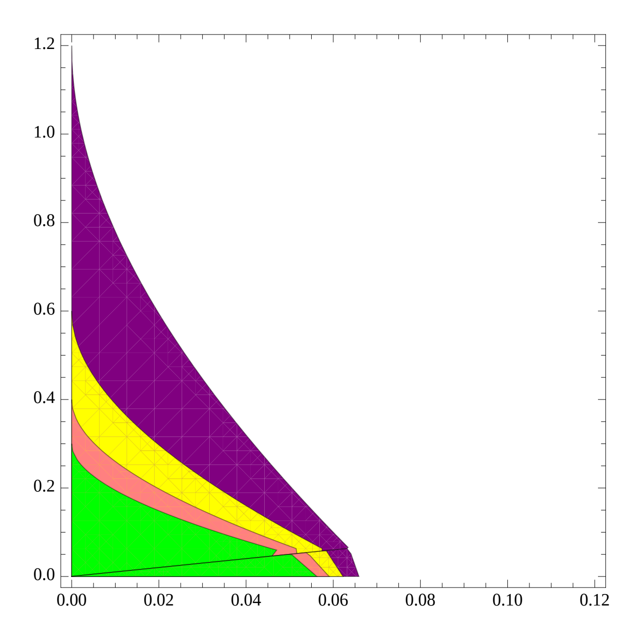


Figura 3.5: Dominios de parámetros de los métodos tipo secante asociados al corolario 3.3 cuando $L\beta = \frac{1}{5}$ y $\lambda = 0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}$ (verde, rosa, amarillo y morado, respectivamente).

i	x_i^*	i	x_i^*	i	x_i^*	i	x_i^*
1	$1.012239\ldots$	3	$1.118079\ldots$	5	$1.159804\ldots$	7	$1.058428\ldots$
2	$1.058428\ldots$	4	$1.159804\ldots$	6	$1.118079\ldots$	8	$1.012239\ldots$

Tabla 3.1: Aproximación de la solución $\mathbf{x}^*$ de (3.14)

3.3.2. Sistema de ecuaciones no lineales no diferenciable

Consideramos ahora la siguiente ecuación integral no lineal de tipo Hammerstein

$$x(s) = \frac{1}{2} + \frac{3}{4}\int_0^1 G(s,t)\,\Big(x(t)^2 + |x(t)|\Big)\,dt, \qquad s \in [0,1], \tag{3.15}$$

donde $x \in C[0,1]$ y G es la función de Green en $[0,1]\times[0,1]$.

A continuación, de nuevo tal y como se hizo en la sección 1.6.1, transformamos (3.15) en el siguiente sistema de ecuaciones no lineales:

$$F(\mathbf{x}) \equiv \mathbf{x} - \mathbf{\frac{1}{2}} - \frac{3}{4}A(\hat{\mathbf{x}} + \tilde{\mathbf{x}}) = 0, \qquad F : \mathbb{R}^8 \longrightarrow \mathbb{R}^8, \tag{3.16}$$

donde

$$\mathbf{x} = (x_1, x_2, \ldots, x_8)^T, \qquad \mathbf{\frac{1}{2}} = \left(\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2}\right)^T, \qquad A = (a_{ij})_{i,j=1}^8,$$
$$\hat{\mathbf{x}} = (x_1^2, x_2^2, \ldots, x_8^2)^T, \qquad \tilde{\mathbf{x}} = (|x_1|, |x_2|, \ldots, |x_8|)^T.$$

En este caso, el operador diferencia dividida considerado es

$$[\mathbf{u}, \mathbf{v}; F] = I - \frac{3}{4}A\,(\mathrm{diag}\{\mathbf{z}\} + \mathrm{diag}\{\mathbf{w}\}),$$

n	$\|\mathbf{x_n} - \mathbf{x}^*\|$	$\|F(\mathbf{x_n})\|$
-1	$2.5980\ldots \times 10^{-1}$	$2.0008\ldots \times 10^{-1}$
0	$1.5980\ldots \times 10^{-1}$	$1.2355\ldots \times 10^{-1}$
1	$4.1222\ldots \times 10^{-3}$	$3.1494\ldots \times 10^{-3}$
2	$4.1117\ldots \times 10^{-5}$	$3.1489\ldots \times 10^{-5}$
3	$1.0346\ldots \times 10^{-8}$	$9.9173\ldots \times 10^{-9}$
4	$2.5597\ldots \times 10^{-14}$	$1.9590\ldots \times 10^{-14}$

Tabla 3.2: Errores absolutos obtenidos con el método tipo secante correspondiente a $\lambda = \frac{1}{2}$ y $\|F(\mathbf{x_n})\|$

donde $\mathbf{z} = (z_1, z_2, \ldots, z_8)^T$ con $z_i = u_i + v_i$, para todo $i = 1, 2, \ldots, 8$, y $\mathbf{w} = (w_1, w_2, \ldots, w_8)^T$ con $w_i = \frac{|u_i| - |v_i|}{u_i - v_i}$, para todo $i = 1, 2, \ldots, 8$.

Eligiendo $\mathbf{x_{-1}} = \left(\frac{2}{5}, \frac{2}{5}, \ldots, \frac{2}{5}\right)^T$ y $\mathbf{x_0} = \left(\frac{1}{2}, \frac{1}{2}, \ldots, \frac{1}{2}\right)^T$ como puntos de salida, $\lambda = \frac{3}{4}$ y la norma del máximo, obtenemos $\alpha = 0.1$, $\beta = 1.2170\ldots$, $\eta = 0.0824\ldots$, $L = 0.1853\ldots$ y $K = 0.0926\ldots$ La ecuación (3.4) se reduce a

$$t\left(1 - \frac{(0.2376\ldots)}{1 - (1.2170\ldots)((0.1863\ldots) + (0.0926\ldots)(2t + 0.0250\ldots))}\right) - (0.0824\ldots) = 0,$$

que tiene dos raíces reales positivas y la más pequeña, denotada por $R = 0.1211\ldots$, cumple la condición (3.5):

$$\beta(L + K(2R + (1-\lambda)\alpha)) = 0.2557\ldots < 1.$$

Así, mediante el teorema 3.2, podemos garantizar la convergencia semilocal del método tipo secante correspondiente a $\lambda = \frac{3}{4}$ a una solución del sistema (3.16). Es más, por el teorema 3.2, la existencia y unicidad de solución está garantizada en la bola $\overline{B(\mathbf{x_0}, 0.1211\ldots)}$.

Después de cuatro iteraciones y usando el criterio de parada $\|\mathbf{x_n} - \mathbf{x_{n-1}}\| < 10^{-16}$, obtenemos la aproximación numérica $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_8^*)^T$ de una solución (3.16) que vemos en la tabla 3.3. En la tabla 3.4 mostramos los errores $\|\mathbf{x_n} - \mathbf{x}^*\|$ obtenidos usando el mismo criterio de parada. Notemos que el vector dado en la tabla 3.3 es una buena aproximación de la solución del sistema (3.16), puesto que $\|F(\mathbf{x}^*)\| \leq$ constante $\times 10^{-16}$. Mostramos la sucesión $\{\|F(\mathbf{x_n})\|\}$ en la tabla 3.4.

i	x_i^*	i	x_i^*	i	x_i^*	i	x_i^*
1	$0.506462\ldots$	3	$0.561738\ldots$	5	$0.583219\ldots$	7	$0.530722\ldots$
2	$0.530722\ldots$	4	$0.583219\ldots$	6	$0.561738\ldots$	8	$0.506462\ldots$

Tabla 3.3: Aproximación de la solución $\mathbf{x}^*$ de (3.16)

n	$\|\mathbf{x_n} - \mathbf{x}^*\|$	$\|F(\mathbf{x_n})\|$
-1	$1.8321\ldots \times 10^{-1}$	$1.5189\ldots \times 10^{-1}$
0	$8.3221\ldots \times 10^{-2}$	$6.9501\ldots \times 10^{-2}$
1	$7.8399\ldots \times 10^{-4}$	$6.5351\ldots \times 10^{-4}$
2	$1.4181\ldots \times 10^{-6}$	$1.1832\ldots \times 10^{-6}$
3	$2.3181\ldots \times 10^{-11}$	$1.9352\ldots \times 10^{-11}$

Tabla 3.4: Errores absolutos obtenidos con el método tipo secante correspondiente a $\lambda = \frac{3}{4}$ y $\|F(\mathbf{x_n})\|$

Parte III

EL MÉTODO DE STEFFENSEN

En esta parte del texto consideraremos un operador $F : \Omega \subset X \to X$, donde Ω es un dominio abierto convexo no vacío contenido en un espacio de Banach X, y nos planteamos la aproximación de una solución de la ecuación

$$F(x) = 0$$

mediante un método iterativo punto a punto que no evalúe derivadas del operador F en cada paso. Este tipo de métodos iterativos tienen gran interés debido a que la derivada del operador F puede ser costosa de evaluar o porque incluso puede no existir, situaciones que son desfavorables para la aplicación del método de Newton, cuyo algoritmo, como bien sabemos, está dado por

$$\begin{cases} \text{dado } x_0 \text{ en } \Omega, \\ x_{n+1} = x_n - [F'(x_n)]^{-1} F(x_n), \quad n \geq 0. \end{cases}$$

Usualmente la no utilización de derivadas pasa por la aproximación de éstas. Así, como bien sabemos, en el caso de espacios de Banach, podemos considerar el concepto de diferencia dividida. Es conocido que, dado $x \in X$, si consideramos un punto $y \in X$ distinto de x, próximo a x, el operador $[x, y; F]$ representa una aproximación de $F'(x)$. Utilizando este hecho, es habitual construir métodos iterativos que utilizan diferencias divididas en vez de derivadas.

Como en la parte anterior de este texto, nuestro punto de partida es el método de Newton. Para eliminar de su algoritmo la evaluación de las derivadas del operador F, utilizamos una aproximación de $F'(x)$ a partir de determinada diferencia dividida. Como el objetivo singular de esta parte del texto es la construcción de un método iterativo punto a punto, teniendo en cuenta las características del operador considerado y tomando como nodos x y $x + F(x)$, podemos utilizar éstos para considerar la aproximación $F'(x) \approx [x, x + F(x); F]$. A partir de esta aproximación, definimos el método de Steffensen, que está dado por el siguiente algoritmo:

$$\begin{cases} \text{dado } x_0 \text{ en } \Omega, \\ x_{n+1} = x_n - [x_n, x_n + F(x_n); F]^{-1} F(x_n), \quad n \geq 0. \end{cases}$$

Evidentemente, el método de Steffensen es un método iterativo punto a punto. Además, por la aproximación de la derivada del operador F considerada en este caso, es conocido que el método de Steffensen tiene convergencia cuadrática [5]. Por otra parte, como resulta evidente, este método iterativo tiene el mismo coste computacional que el método de Newton. Por tanto, la eficiencia computacional ([35], [45]) del método de Steffensen es la misma que la del método de Newton. Pese a las buenas características que tiene el método de Steffensen, misma eficiencia computacional que el método de Newton y la no evaluación de derivadas, es mucho menos utilizado que el método de Newton para aproximar soluciones de la ecuación $F(x) = 0$. Nuestro objetivo principal se va a centrar entonces en mejorar la aplicabilidad del método de Steffensen.

Como hemos dicho en la sección 1.5 del capítulo 1, el aspecto central a la hora de estudiar la aplicabilidad de un método iterativo es su accesibilidad. Experimentalmente, podemos verla mediante las cuencas de atracción, que recordamos que son el conjunto de todos los puntos de salida a partir de los cuales el método iterativo converge a una solución de $F(x) = 0$, fijada una tolerancia o un número máximo de iteraciones.

A continuación, en las figuras 3.6 y 3.7 mostramos las cuencas de atracción asociadas a las tres raíces $z^* = 1$, $z^{**} = \exp\left(\frac{2\pi i}{3}\right)$ y $z^{***} = \exp\left(\frac{-2\pi i}{3}\right)$ de la ecuación $F(z) = z^3 - 1 = 0$,

donde $F : \mathbb{C} \to \mathbb{C}$, cuando aplicamos respectivamente los métodos de Newton y de Steffensen. Para representar las cuencas de atracción de las tres raíces de la ecuación anterior, hemos seguido la misma estrategia que la descrita en la sección 1.5.2 para el método de Newton.

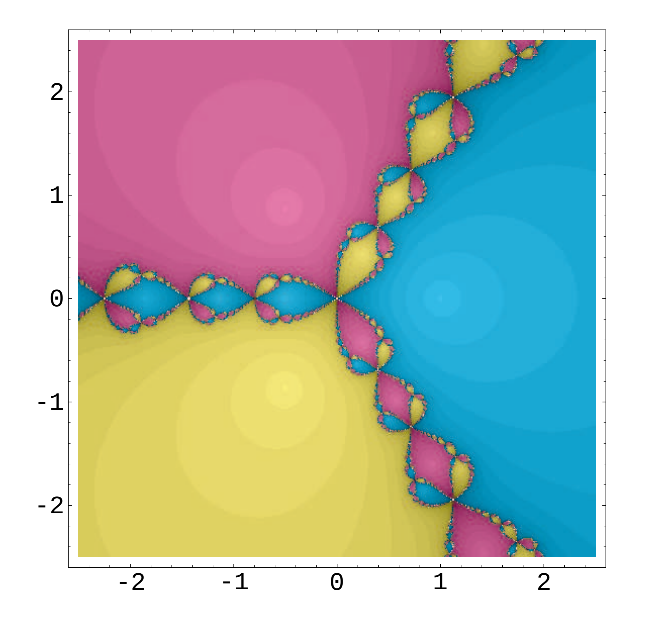


Figura 3.6: Cuencas de atracción asociadas a las tres raíces de $F(z) = z^3 - 1 = 0$ para el método de Newton.

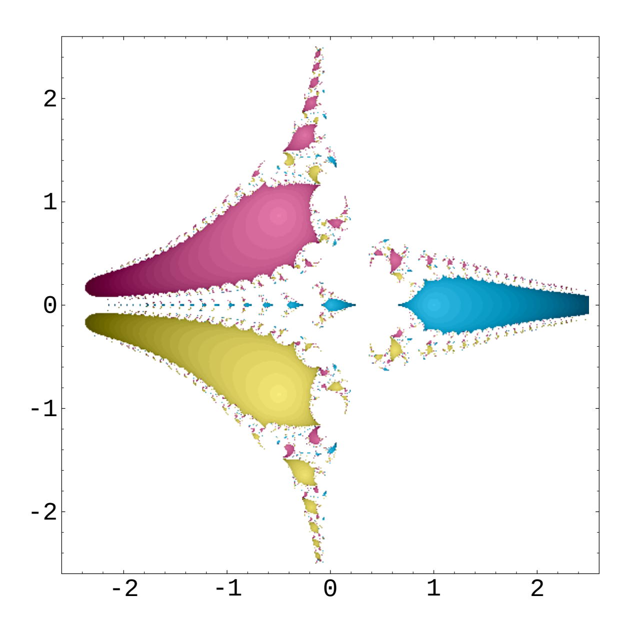


Figura 3.7: Cuencas de atracción asociadas a las tres raíces de $F(z) = z^3 - 1 = 0$ para el método de Steffensen.

Vemos que las cuencas de atracción, cuando aplicamos el método de Steffensen, se reducen considerablemente con respecto a las del método de Newton, lo que justifica claramente su menor uso con respecto al de Newton a la hora de resolver ecuaciones no lineales.

También hemos indicado en la sección 1.5 del capítulo 1 que, aparte del estudio empírico de la accesibilidad de un método iterativo dado por las cuencas de atracción, podemos realizar un estudio de la accesibilidad del método iterativo a partir de las condiciones de convergencia impuestas al método iterativo, mediante la región de accesibilidad y el dominio de parámetros.

En el capítulo 4 de esta parte del texto realizamos un estudio de la convergencia semilocal del método de Steffensen mediante el principio de la mayorante de Kantorovich, de manera que podamos comparar su accesibilidad con la del método de Newton bajo condiciones de tipo Kantorovich. Así, vemos que el dominio de parámetros asociado al resultado de convergencia semilocal del método de Steffensen es más pequeño que el del método de Newton. Para solventar esta situación, construimos un método iterativo híbrido (predictor-corrector) que no evalúe derivadas en cada paso y con la misma accesibilidad que el método de Newton.

La utilización de la teoría de Kantorovich en el capítulo 4 para estudiar la convergencia semilocal de los métodos iterativos nos obliga a considerar ecuaciones con operadores diferenciables. Así, en el capítulo 5, modificamos las condiciones de convergencia consideradas en el capítulo 4 y desarrollamos una teoría, basada en relaciones de recurrencia, que proporciona resultados de convergencia semilocal y permite aproximar soluciones de ecuaciones definidas mediante operadores cualesquiera, tanto diferenciables como no diferenciables. Después, a partir del estudio de los dominios de parámetros asociados a los nuevos resultados, construimos un método iterativo híbrido (predictor-corrector) aplicable a ecuaciones con operadores cualesquiera y con una buena accesibilidad de soluciones.

Capítulo 4

Situación diferenciable

Comenzamos este capítulo realizando un estudio de la convergencia semilocal del método de Steffensen mediante el principio de la mayorante de Kantorovich. Hemos visto antes que para establecer un resultado de convergencia semilocal se exigen dos tipos de condiciones: condiciones sobre los puntos de salida y condiciones sobre el operador implicado. Como nuestro primer interés se centra en el estudio comparativo de la accesibilidad de los métodos de Newton y de Steffensen a partir de la teoría de Kantorovich, exigimos la misma condición al operador F:

$$\|F'(x) - F'(y)\| \leq C\,\|x - y\|, \quad C \geq 0, \quad \forall x, y \in \Omega;$$

es decir, que F' sea Lipschitz continua en Ω. Notemos en primer lugar que, a partir de la condición anterior, siempre existe la diferencia dividida

$$[x, y; F] = \int_0^1 F'\left(x + \tau(x - y)\right) d\tau,$$

para cualesquiera que sean los puntos $x, y \in \Omega$. Si $x = y$, entonces $[x, x; F] = F'(x)$, [37]. Luego el método de Steffensen está bien definido en cada paso.

A partir del estudio de la convergencia semilocal del método de Steffensen, que realizaremos mediante la teoría de Kantorovich, probamos que la accesibilidad del método de Newton es mucho mejor que la que se obtiene para el método de Steffensen en las mismas condiciones.

En la sección 4.1 obtenemos un teorema de convergencia semilocal para el método de Steffensen mediante la teoría de Kantorovich. Luego, comparamos la accesibilidad de los métodos de Newton y de Steffensen. Adelantamos que las diferencias que se observan experimentalmente con las cuencas de atracción de las tres raíces de la ecuación $F(z) = z^3 - 1 = 0$ se confirman en el estudio teórico.

A continuación, en la sección 4.2, consideramos el método simplificado de Newton, cuyo algoritmo es

$$\begin{cases} \text{dado } z_0 \text{ en } \Omega, \\ z_{n+1} = z_n - [F'(z_0)]^{-1} F(z_n), \quad n \geq 0, \end{cases} \tag{4.1}$$

observamos que no se evalúa F' en un nuevo punto en cada paso, sino que basta considerar un punto z_0 de Ω donde se evalúe F' y vemos que este método tiene la misma accesibilidad que el método de Newton.

Teniendo en cuenta lo anterior en la sección 4.3, construimos un método iterativo híbrido (predictor-corrector) considerando el método simplificado de Newton (predictor) en unos primeros pasos para localizar un punto de salida favorable para el método de Steffensen

(método corrector). Este método híbrido tiene la misma accesibilidad que el método de Newton, pero no evalúa derivadas en cada paso. En particular, saliendo desde los mismos puntos de salida que el método simplificado de Newton, garantizamos la convergencia del método híbrido. Finalmente, en la sección 4.4, veremos una aplicación en la que consideramos un sistema de ecuaciones no lineales que ilustra todo lo anterior.

4.1. Método corrector: el método de Steffensen

4.1.1. Convergencia semilocal

A continuación, suponemos que se satisfacen las siguientes condiciones:

(A1) $\|F(x_0)\| \leq \delta$,

(A2) existe $\Gamma_0 = [F'(x_0)]^{-1} \in \mathcal{L}(X, X)$, para $x_0 \in \Omega$, y es tal que $\|\Gamma_0\| \leq \theta$,

(A3) $\|F'(x) - F'(y)\| \leq C\|x - y\|,\ C \geq 0,\ x, y \in \Omega$.

En primer lugar, establecemos la convergencia semilocal del método de Steffensen, bajo las condiciones (A1)–(A3), de manera similar a la dada para el método de Newton en el teorema de Newton-Kantorovich (teorema 1.53 del capítulo 1), utilizando entonces el principio de la mayorante de Kantorovich.

Lo primero que vamos a hacer es probar la existencia del operador $[x_0, y_0; F]^{-1} \in \mathcal{L}(X, X)$ para $x_0, y_0 \in \Omega$. Así,

$$\begin{aligned} \|I - \Gamma_0[x_0, y_0; F]\| &\leq \|\Gamma_0\|\|F'(x_0) - [x_0, y_0; F]\| \\ &= \|\Gamma_0\| \left\| F'(x_0) - \int_0^1 F'(x_0 + \tau(y_0 - x_0))\, d\tau \right\| \\ &\leq \|\Gamma_0\| \int_0^1 \|F'(x_0 + \tau(y_0 - x_0)) - F'(x_0)\|\, d\tau \\ &\leq \frac{1}{2} C\delta\theta, \end{aligned}$$

de modo que, si $C\delta\theta < 2$, existe $[x_0, y_0; F]^{-1} \in \mathcal{L}(X, X)$, para $x_0, y_0 \in \Omega$, y es tal que

$$\left\| [x_0, y_0; F]^{-1} \right\| \leq \frac{2\theta}{2 - C\delta\theta} = b.$$

A continuación, definimos el polinomio cuadrático

$$q(s) = \frac{\ell}{2}s^2 - \frac{s}{b} + \delta, \qquad \ell = C\left(1 + \frac{1}{b}\right), \qquad s \in [0, s'], \tag{4.2}$$

tal que $s^* \leq s^{**} < s'$, donde denotamos la raíz positiva más pequeña del polinomio q por $s^* = \dfrac{1 - \sqrt{1 - 2\ell\delta b^2}}{\ell b}$ y la mayor por $s^{**} = \dfrac{1 + \sqrt{1 - 2\ell\delta b^2}}{\ell b}$, y la sucesión $\{s_n\}$ definida por

$$\left\{ \begin{array}{l} s_0 = 0, \\ s_{n+1} = s_n - \dfrac{q(s_n)}{q'(s_n)}, \quad n \geq 0. \end{array} \right. \tag{4.3}$$

Observemos que la sucesión $\{s_n\}$ converge de forma creciente a s^*, sin más que tener en cuenta que $\{s_n\}$ es una sucesión monótona no decreciente y está acotada superiormente por s^*.

Teorema 4.1. *Sean X un espacio de Banach y $F : \Omega \subset X \to X$ un operador una vez continuamente diferenciable Fréchet definido en un dominio abierto convexo no vacío Ω. Si se cumplen las condiciones (A1)–(A3), se satisfacen las condiciones*

$$C\delta\theta \leq 2, \qquad \ell\delta b^2 \leq \frac{1}{2} \tag{4.4}$$

y $B(x_0, s^* + \delta) \subset \Omega$, entonces el método de Steffensen converge, a partir de x_0, a una solución x^* de $F(x) = 0$. Además, la solución x^* y las iteraciones x_n pertenecen a $\overline{B(x_0, s^* + \delta)}$ y x^* es única en $B(x_0, s^{**} + \delta) \cap \Omega$, donde $s^{**} = \frac{1+\sqrt{1-2\ell\delta b^2}}{\ell b}$.

Demostración. Probamos la convergencia semilocal del método de Steffensen utilizando un proceso inductivo.

Notemos que

$$\|x_1 - x_0\| = \left\|[x_0, y_0; F]^{-1} F(x_0)\right\| \leq b\delta = s_1 - s_0 < s^* + \delta,$$

de manera que $x_1 \in B(x_0, s^* + \delta) \subset \Omega$ y podamos definir x_2. A continuación, como

$$\begin{aligned}
F(x_1) &= F(x_1) - F(x_0) - [x_0, y_0; F](x_1 - x_0) \\
&= \int_{x_0}^{x_1} (F'(z) - F'(x_0))\, dz + (F'(x_0) - [x_0, y_0; F])\,(x_1 - x_0) \\
&= \int_0^1 (F'(x_0 + \tau(x_1 - x_0)) - F'(x_0))\, d\tau (x_1 - x_0) + (F'(x_0) - [x_0, y_0; F])\,(x_1 - x_0),
\end{aligned}$$

vemos que

$$\begin{aligned}
\|F(x_1)\| &\leq \frac{C}{2}\|x_1 - x_0\|^2 + \frac{C}{2}\|F(x_0)\|\|x_1 - x_0\| \\
&\leq \frac{C}{2}(s_1 - s_0)^2 + \frac{C}{2}q(s_0)(s_1 - s_0) \\
&= \frac{C}{2}\,(1 - q'(s_0))\,(s_1 - s_0)^2 \\
&= \frac{\ell}{2}(s_1 - s_0)^2 \\
&= q(s_1).
\end{aligned}$$

Ahora, puesto que la sucesión $\{s_n\}$, dada en (4.3), es creciente y el polinomio (4.2) es decreciente en $[0, s^*]$, se tiene

$$\|y_1 - x_0\| \leq \|x_1 - x_0\| + \|F(x_1)\| \leq s_1 - s_0 + q(s_1) < s^* + q(s_0) = s^* + \delta.$$

Notemos que, para poder definir x_2, necesitamos la existencia del operador $[x_1, y_1; F]^{-1}$. Teniendo de nuevo en cuenta nuevamente que la sucesión (4.3) es creciente y el polinomio

(4.2) es decreciente en $[0, s^*]$, se tiene

$$\begin{aligned}
\|I - \Gamma_0[x_1, y_1; F]\| &\leq \|\Gamma_0\| \, \|F'(x_0) - [x_1, y_1; F]\| \\
&= \|\Gamma_0\| \left(\|F'(x_0) - F'(x_1)\| + \|F'(x_1) - [x_1, y_1; F]\|\right) \\
&\leq \|\Gamma_0\| \left(C \, \|x_1 - x_0\| + \int_0^1 \|F'(x_1 + \tau(y_1 - x_1)) - F'(x_1)\| \, d\tau \right) \\
&\leq \|\Gamma_0\| \left(C \, \|x_1 - x_0\| + \frac{C}{2}\|F(x_1)\| \right) \\
&\leq \|\Gamma_0\| \left(C \, (s_1 - s_0) + \frac{C}{2} q(s_1) \right) \\
&\leq \|\Gamma_0\| \left(C s_1 + \frac{C}{2} q(s_0) \right) \\
&\leq \|\Gamma_0\| \, C \left(1 - \frac{q'(s_0)}{2} \right) s_1 \\
&\leq \theta \left(q'(s_1) + \frac{1}{b} \right) \\
&\leq \theta \left(q'(s_1) + \frac{1}{\theta} \right) \\
&< 1,
\end{aligned}$$

de modo que el operador $[x_1, y_1; F]^{-1}$ existe y es tal que

$$\|[x_1, y_1; F]^{-1}\| \leq \frac{\theta}{1 - \|I - \Gamma_0[x_1, y_1; F]\|} \leq -\frac{1}{q'(s_1)}.$$

Entonces,

$$\|x_2 - x_1\| \leq -\frac{q(s_1)}{q'(s_1)} \leq s_2 - s_1,$$

$$\|x_2 - x_0\| \leq \|x_2 - x_1\| + \|x_1 - x_0\| \leq s_2 - s_0 < s^* < s^* + \delta$$

y $x_2 \in B(x_0, s^* + \delta) \subset \Omega$, lo que permite definir y_2.

Ahora, a partir de

$$\begin{aligned}
F(x_n) &= F(x_n) - F(x_{n-1}) - [x_{n-1}, y_{n-1}; F](x_n - x_{n-1}) \\
&= \int_{x_{n-1}}^{x_n} (F'(z) - F'(x_{n-1})) \, dz + (F'(x_{n-1}) - [x_{n-1}, y_{n-1}; F]) \, (x_n - x_{n-1}) \\
&= \int_0^1 (F'(x_{n-1} + \tau(x_n - x_{n-1})) - F'(x_{n-1})) \, d\tau (x_n - x_{n-1}) \\
&\quad + (F'(x_{n-1}) - [x_{n-1}, y_{n-1}; F]) \, (x_n - x_{n-1}) \\
&= \int_0^1 (F'(x_{n-1} + \tau(x_n - x_{n-1})) - F'(x_{n-1})) \, d\tau (x_n - x_{n-1}) \\
&\quad + \int_0^1 (F'(x_{n-1}) - F'(x_{n-1} + \tau(y_{n-1} - x_{n-1}) \, d\tau) \, (x_n - x_{n-1})
\end{aligned}$$

y

$$\begin{aligned}
q(s_n) &= q(s_n) - q(s_{n-1}) - q'(s_{n-1})(s_n - s_{n-1}) \\
&= \int_0^1 (q'(s_{n-1} + \tau(s_n - s_{n-1})) - q'(s_{n-1}))\, d\tau (s_n - s_{n-1}) \\
&= \ell \int_0^1 \tau (s_n - s_{n-1})^2 \, d\tau \\
&= \frac{\ell}{2}(s_n - s_{n-1})^2,
\end{aligned}$$

se tiene

$$\|F(x_n)\| \leq q(s_n), \quad \forall n \in \mathbb{N},$$

puesto que

$$\begin{aligned}
\|F(x_n)\| &\leq \int_0^1 C\tau \|x_n - x_{n-1}\|^2 \, d\tau + \int_0^1 C\tau \|F(x_{n-1})\| \|x_n - x_{n-1}\| \, d\tau \\
&\leq \frac{C}{2}(s_n - s_{n-1})^2 + \frac{C}{2} q(s_{n-1})(s_n - s_{n-1}) \\
&= \frac{C}{2} \left(1 - q'(s_{n-1})\right) (s_n - s_{n-1})^2 \\
&\leq \frac{\ell}{2}(s_n - s_{n-1})^2 \\
&= q(s_n).
\end{aligned}$$

Entonces, como el polinomio q es decreciente en $[0, s^*]$, se sigue

$$\|y_n - x_0\| \leq \|x_n - x_0\| + \|F(x_n)\| \leq s_n + q(s_n) < s^* + q(s_0) < s^* + \delta$$

y, por tanto, $y_n \in B(x_0, s^* + \delta) \subset \Omega$.

A continuación, probamos la existencia del operador $[x_n, y_n; F]^{-1}$. Como

$$\begin{aligned}
\|I - \Gamma_0[x_n, y_n; F]\| &\leq \|\Gamma_0\| \|F'(x_0) - [x_n, y_n; F]\| \\
&= \|\Gamma_0\| \left(\|F'(x_0) - F'(x_n)\| + \|F'(x_n) - [x_n, y_n; F]\|\right) \\
&\leq \|\Gamma_0\| \left(C\|x_n - x_0\| + \int_0^1 \|F'(x_n + \tau(y_n - x_n)) - F'(x_n)\| \, d\tau \right) \\
&\leq \|\Gamma_0\| \left(C\|x_n - x_0\| + \frac{C}{2}\|F(x_n)\| \right) \\
&\leq \|\Gamma_0\| \left(C(s_n - s_0) + \frac{C}{2} q(s_n) \right) \\
&\leq \|\Gamma_0\| \left(Cs_n - \frac{C}{2} q'(s_n)(s_n - s_{n-1}) \right) \\
&\leq \|\Gamma_0\| C \left(1 - \frac{q'(s_0)}{2} \right) s_n \\
&\leq \theta \left(q'(s_n) + \frac{1}{b} \right) \\
&\leq \theta \left(q'(s_n) + \frac{1}{\theta} \right) \\
&< 1,
\end{aligned}$$

tenemos que el operador $[x_n, y_n; F]^{-1}$ existe y es tal que

$$\|[x_n, y_n; F]^{-1}\| \leq \frac{\theta}{1 - \|I - \Gamma_0[x_n, y_n; F]\|} \leq -\frac{1}{q'(s_n)}.$$

Por tanto,

$$\|x_{n+1} - x_n\| \leq \|[x_n, y_n; F]^{-1}\|\|F(x_n)\| \leq -\frac{q(s_n)}{q'(s_n)} = s_{n+1} - s_n,$$

$$\|x_{n+1} - x_0\| \leq \|x_{n+1} - x_n\| + \|x_n - x_0\| \leq s_{n+1} - s_0 < s^* < s^* + \delta,$$

y, como $\{s_n\}$ es una sucesión de Cauchy, se deduce que $\{x_n\}$ también lo es y, en consecuencia, $\{x_n\}$ converge a $x^* \in \overline{B(x_0, s^* + \delta)}$. Para ver que x^* es una solución de la ecuación $F(x) = 0$, basta tener en cuenta que $\|F(x_n)\| \leq q(s_n)$ y, por la continuidad de F, tenemos $F(x^*) = 0$.

A continuación, probamos la unicidad de solución. Suponemos que tenemos otra solución $z^* \in B(x_0, s^{**} + \delta) \cap \Omega$ de la ecuación $F(x) = 0$ distinta de x^*. Consideramos

$$F(z^*) - F(x^*) = \int_{x^*}^{z^*} F'(x)\,dx = \int_0^1 F'(x^* + t(z^* - x^*))(z^* - x^*)\,dt = 0$$

y el operador $J = \displaystyle\int_0^1 F'(x^* + t(z^* - x^*))\,dt$. Teniendo en cuenta que

$$\begin{aligned}
\|I - \Gamma_0 J\| &\leq \|\Gamma_0\|\|F'(x_0) - J\| \\
&\leq \|\Gamma_0\| \int_0^1 \|F'(x^* + t(z^* - x^*)) - F'(x_0)\|\,dt \\
&\leq \theta C \int_0^1 \left((1 - t)\|x^* - x_0\| + t(\|z^* - x_0\|)\right)\,dt \\
&< \frac{\theta C}{2}\left(\frac{2}{\ell b} + \delta\right) \\
&< 1,
\end{aligned}$$

puesto que $C\delta\theta \leq 2$, el operador J es inversible y, por tanto, $x^* = z^*$. ■

4.1.2. Accesibilidad

Buscando cierto paralelismo con las cuencas de atracción vistas en las figuras 3.6 y 3.7, consideramos ahora la otra forma experimental de estudiar la accesibilidad de un método iterativo: las regiones de accesibilidad.

En la figura 4.1, se muestran las regiones de accesibilidad de la solución $z = 1$ de la ecuación compleja $F(z) = z^3 - 1 = 0$, una vez fijado el dominio complejo $[0.8, 1.6] \times [-0.2, 0.2]$ que conduce a que $C = 6|1.6 + 0.2i|$, para los métodos de Newton, según el teorema de Newton-Kantorovich, y de Steffensen, según el teorema 4.1. Notamos que las regiones están superpuestas. Observamos entonces la gran diferencia que existe entre las dos regiones de accesibilidad, tal y como ocurría con las cuencas de atracción, volviéndose a poner de manifiesto la reducida accesibilidad del método de Steffensen.

Recordemos, sección 1.5.2 del capítulo 1, que para representar gráficamente la región de accesibilidad para el método de Newton, coloreamos los puntos x_0 que verifican la condición $C\delta\theta^2 \leq \frac{1}{2}$ del teorema de Newton-Kantorovich. Para representar gráficamente la región de

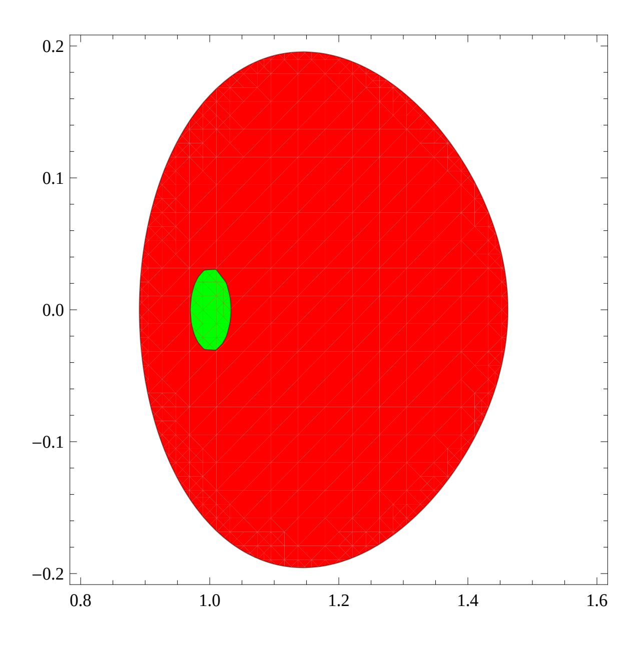


Figura 4.1: Regiones de accesibilidad de la solución $z = 1$ de la ecuación compleja $F(z) = z^3 - 1 = 0$ para los métodos de Newton (región roja) y de Steffensen (región verde), según los teoremas de Newton-Kantorovich y 4.1, respectivamente.

accesibilidad para el método de Steffensen, coloreamos los puntos x_0 que verifican las dos condiciones dadas en (4.4) del teorema 4.1.

Por otra parte, si comparamos teóricamente la accesibilidad de ambos métodos mediante los dominios de parámetros asociados a los teoremas de Newton-Kantorovich y 4.1, vemos claramente en la figura 4.2 que la accesibilidad del método de Steffensen se reduce considerablemente con respecto a la del método de Newton. Notamos que las regiones están superpuestas.

Recordemos, sección 1.5.2 del capítulo 1, que para representar gráficamente el dominio de parámetros del método de Newton asociado al teorema de Newton-Kantorovich, coloreamos en el plano xy el conjunto de puntos $\{(x, y) \in \mathbb{R}^2 : \; x^2 y \leq \frac{1}{2}\}$, donde $x = \theta$ e $y = C\delta$. Para representar gráficamente el dominio de parámetros del método de Steffensen asociado al teorema 4.1, coloreamos en el mismo plano xy el conjunto de puntos que cumplen las dos condiciones dadas en (4.4), véase la figura 4.2.

4.2. Método predictor: el método simplificado de Newton

Con el objetivo de aumentar la aplicabilidad del método de Steffensen, vamos a construir, como se ha indicado en la introducción, un método iterativo híbrido (predictor-corrector) que nos permita localizar puntos de salida para el método de Steffensen y que, a partir de ellos, esté garantizada la convergencia del método. Para ello, tenemos que considerar un método predictor apropiado. Así, consideramos como método predictor el método simplificado de Newton (4.1), que tiene la misma accesibilidad que el método de Newton, pero sin la necesidad de tener que evaluar derivadas en cada paso de iteración. Como el operador F es derivable,

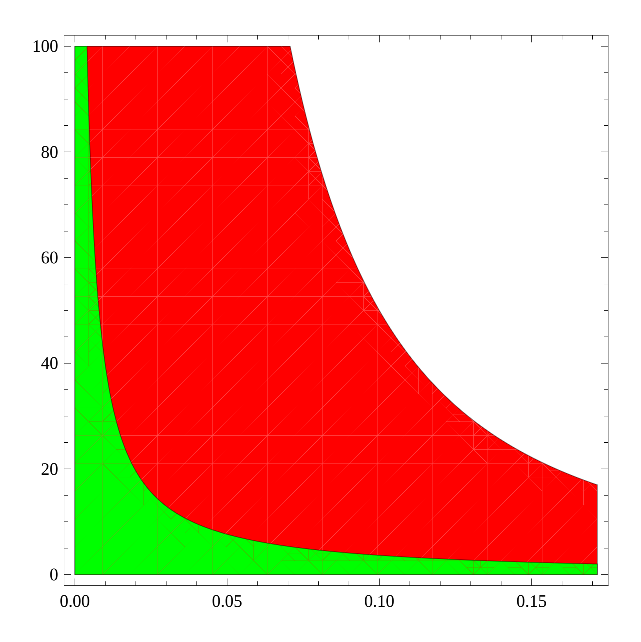


Figura 4.2: Dominios de parámetros de los métodos de Newton (región roja) y de Steffensen (región verde) asociados a los teoremas de Newton-Kantorovich y 4.1, respectivamente.

dado el punto de salida x_0, existirá el operador $F'(x_0)$, que es la única derivada que se evalúa a la hora de aplicar el método simplificado de Newton (4.1).

4.2.1. Convergencia semilocal

Acabamos de ver que la accesibilidad del método de Steffensen es peor que la del método de Newton. Sin embargo, como veremos a continuación, la accesibilidad del método simplificado de Newton (4.1) es la misma que la del método de Newton, si bien tiene menor velocidad de convergencia, puesto que su R-orden de convergencia es al menos lineal. Esto nos lleva a considerar un método iterativo híbrido (predictor-corrector) formado por los métodos simplificado de Newton (4.1) y de Steffensen. Así, mediante un método predictor con buena accesibilidad y un método corrector con buena velocidad de convergencia, obtendremos un método híbrido con características similares a las del método de Newton.

Comenzamos viendo que el método simplificado de Newton (4.1) tiene la misma accesibilidad que el método de Newton. Establecemos entonces bajo qué condiciones para el punto de salida z_0 y el operador F podemos garantizar la convergencia del método simplificado de Newton (4.1) a una solución de la ecuación $F(x) = 0$. Para ello, suponemos que se cumplen las siguientes condiciones:

(H1) $\|F(z_0)\| \leq \delta_0$,

(H2) existe $\overline{\Gamma}_0 = [F'(z_0)]^{-1} \in \mathcal{L}(X, X)$, para $z_0 \in \Omega$, y es tal que $\|\overline{\Gamma}_0\| \leq \theta_0$,

(H3) $\|F'(x) - F'(y)\| \leq C\|x - y\|$, $\quad C \geq 0$, $\quad x, y \in \Omega$.

A partir de (H1)–(H3) definimos el polinomio cuadrático

$$p(t) = \frac{C}{2}t^2 - \frac{t}{\theta_0} + \delta_0, \quad t \in [t_0, t'], \tag{4.5}$$

donde $t^* \leq t^{**} \leq t'$, y denotamos por t^* y t^{**} sus dos raíces positivas. Definimos también la función

$$h(t) = t + \theta_0 p(t) \tag{4.6}$$

y consideramos la sucesión $\{t_n\}$ dada por

$$\begin{cases} t_0 = 0, \\ t_{n+1} = h(t_n) = t_n + \theta_0 p(t_n), \quad n \geq 0. \end{cases} \tag{4.7}$$

Lema 4.2. *La sucesión (4.7) converge de forma creciente a t^*.*

Demostración. De (4.7), se sigue

$$t_1 - t_0 = \theta_0 p(t_0) > 0,$$

$$t_1 - t^* = h(t_0) - h(t^*) = h'(\theta_0)(t_0 - t^*) < 0,$$

de modo que $t_1 > t_0$ y $t_1 < t^*$. Supongamos que $t_k \geq t_{k-1}$ y $t_k < t^*$, para $k = 1, 2 \ldots, n-1$. Entonces, como $t_{n-1} < t^*$, se tiene $p(t_{n-1}) \geq 0$ y

$$t_n - t_{n-1} = \theta_0 p(t_{n-1}) \geq 0.$$

Además,

$$t_n - t^* = h(t_{n-1}) - h(t^*) = h'(\theta_{n-1})(t_{n-1} - t^*) < 0.$$

Por tanto, la sucesión $\{t_n\}$ es creciente y $t_n < t^*$, para todo $n \in \mathbb{N}$. Luego, existe $\tilde{t} = \lim_n t_n$. Ahora bien, como

$$\tilde{t} = \lim_n t_{n+1} = \lim_n \left(t_n + \theta_0 p(t_n)\right),$$

tenemos $\theta_0 p(\tilde{t}) = 0$, de donde se sigue $\tilde{t} = t^*$. ■

Teorema 4.3. *Sean X un espacio de Banach y $F : \Omega \subset X \to X$ un operador una vez continuamente diferenciable Fréchet definido en un dominio abierto convexo no vacío Ω. Supongamos que se verifican las condiciones (H1)–(H3) y que se satisfacen*

$$C\delta_0\theta_0^2 \leq \frac{1}{2} \tag{4.8}$$

y $B(z_0, t^*) \subset \Omega$, donde $t^* = \frac{1-\sqrt{1-2C\delta_0\theta_0^2}}{C\theta_0}$. Entonces, la ecuación $F(x) = 0$ tiene una solución z^* y el método simplificado de Newton (4.1) converge a z^* empezando en z_0. Además, $z_n, z^* \in \overline{B(z_0, t^*)}$, para todo $n \in \mathbb{N}$. Por otra parte, si $C\delta_0\theta_0^2 < \frac{1}{2}$, z^* es la única solución de $F(x) = 0$ en $B(x_0, t^{**}) \cap \Omega$, donde $t^{**} = \frac{1+\sqrt{1-2C\delta_0\theta_0^2}}{C\theta_0}$.

Demostración. En primer lugar, consideramos $z_1 = z_0 - \overline{\Gamma}_0 F(z_0)$. En este caso,

$$\|z_1 - z_0\| = \|\overline{\Gamma}_0 F(z_0)\| \leq \theta_0\delta_0 < t_1 - t_0 < t^*.$$

Luego, $z_1 \in B(z_0, t^*) \subset \Omega$ y podemos definir $z_2 = z_1 - \overline{\Gamma}_0 F(z_1)$.

Ahora,

$$\begin{aligned} F(z_1) &= F(z_1) - F(z_0) - F'(z_0)(z_1 - z_0) \\ &= \int_{z_0}^{z_1} (F'(u) - F'(z_0))\ du \\ &= \int_0^1 (F'(z_0 + \tau(z_1 - z_0)) - F'(z_0))\ d\tau(z_1 - z_0) \end{aligned}$$

y

$$\|F(z_1)\| \leq \frac{C}{2}\|z_1 - z_0\|^2 \leq \frac{C}{2}(t_1 - t_0)^2,$$

de forma que

$$\|z_2 - z_1\| \leq \frac{\theta_0 C}{2} t_1^2.$$

Teniendo en cuenta que $t_2 - t_1 = \theta_0 p(t_1) = \dfrac{\theta_0 C t_1^2}{2}$, se sigue

$$\begin{aligned} \|z_2 - z_1\| &\leq t_2 - t_1, \\ \|z_2 - z_0\| &\leq \|z_2 - z_1\| + \|z_1 - z_0\| \leq t_2 - t_0 < t^*. \end{aligned}$$

Luego, $z_2 \in B(z_0, t^*) \subset \Omega$ y podemos definir z_3.

Ahora, a partir de

$$\begin{aligned} F(z_n) &= F(z_n) - F(z_{n-1}) - F'(z_0)(z_n - z_{n-1}) \\ &= \int_{z_{n-1}}^{z_n} (F'(u) - F'(z_0))\ du \\ &= \int_0^1 (F'(z_{n-1} + \tau(z_n - z_{n-1})) - F'(z_0))\ d\tau(z_n - z_{n-1}), \end{aligned}$$

y

$$\begin{aligned} p(t_n) &= p(t_n) - p(t_{n-1}) + \frac{1}{\theta_0}(t_n - t_{n-1}) \\ &= \int_{t_{n-1}}^{t_n} (p'(u) - p'(t_0))\ du \\ &= \int_0^1 (p'(t_{n-1} + \tau(t_n - t_{n-1})) - p'(t_0))\ d\tau(t_n - t_{n-1}) \\ &= C\left((t_{n-1} - t_0) + \frac{1}{2}(t_n - t_{n-1})\right)(t_n - t_{n-1}), \end{aligned}$$

se tiene

$$\|F(z_n)\| \leq p(t_n), \quad \forall n \in \mathbb{N},$$

puesto que

$$\begin{aligned} \|F(z_n)\| &\leq \int_0^1 C\left(\|z_{n-1} - t_0\| + \tau\|z_n - z_{n-1}\|\right)\ d\tau\|z_n - z_{n-1}\| \\ &\leq C\left((t_{n-1} - t_0) + \frac{1}{2}(t_n - t_{n-1})\right)(t_n - t_{n-1}). \end{aligned}$$

Entonces,

$$\|z_n - z_{n-1}\| \leq \theta_0 p(t_n) = t_n - t_{n-1}$$

y, por tanto, como la sucesión $\{t_n\}$ es de Cauchy, se deduce que la sucesión $\{z_n\}$ también lo es y, en consecuencia, $\{z_n\}$ converge a $z^* \in \overline{B(z_0, t^*)}$. Veamos que z^* es solución de la ecuación $F(x) = 0$. Hemos probado que $\|F(z_n)\| \leq p(t_n)$, de manera que $F(z^*) = 0$ se sigue por continuidad.

Finalmente, suponemos que existe otra solución $y^* \in B(z_0, t^{**}) \cap \Omega$ de la ecuación $F(x) = 0$ distinta de z^*. Consideramos

$$F(y^*) - F(z^*) = \int_{z^*}^{y^*} F'(x)\,dx = \int_0^1 F'(z^* + t(y^* - z^*))(y^* - z^*)\,dt = 0$$

y el operador $J = \displaystyle\int_0^1 F'(z^* + t(y^* - z^*))\,dt$. Teniendo en cuenta que

$$\begin{aligned}
\|I - \overline{\Gamma}_0 J\| &\leq \|\overline{\Gamma}_0\|\|F'(z_0) - J\| \\
&\leq \|\overline{\Gamma}_0\| \int_0^1 \|F'(z^* + t(y^* - z^*)) - F'(z_0)\|\,dt \\
&\leq \theta_0 C \int_0^1 \left((1-t)\|z^* - z_0\| + t\|y^* - z_0\|\right)\,dt \\
&< \frac{C\theta_0}{2}(t^{**} + t^*) \\
&= 1,
\end{aligned}$$

vemos, por el lema de Banach (lema 1.22), que el operador J es inversible y, por tanto, $x^* = z^*$. ■

4.2.2. Accesibilidad

En la figura 4.3 se muestran las regiones de accesibilidad de la raíz $z = 1$ de la ecuación compleja $F(z) = z^3 - 1 = 0$, una vez fijado el dominio complejo $[0.8, 1.6] \times [-0.2, 0.2]$ que conduce a que $C = 6|1.6 + 0.2i|$, para los métodos simplificado de Newton (4.1), según el teorema 4.3, y de Steffensen, según el teorema 4.1. Notamos que las regiones están superpuestas. Observamos entonces que ocurre lo mismo que cuando comparábamos las regiones de accesibilidad de los métodos de Newton y de Steffensen: la gran diferencia entre ambas regiones. De hecho, las regiones de accesibilidad para los métodos de Newton y simplificado de Newton (4.1) son las mismas. Esto nos indica que el dominio de puntos de salida para obtener una aproximación de la raíz $z = 1$ de la ecuación compleja $F(z) = z^3 - 1 = 0$ es mucho mayor para el método simplificado de Newton (4.1) que para el de Steffensen.

Por otra parte, si representamos gráficamente el dominio de parámetros del método simplificado de Newton (4.1) asociado al teorema 4.3, figura 4.4, vemos, como es obvio, que es el mismo que el del método de Newton asociado al teorema de Newton-Kantorovich, puesto que las condiciones de convergencia impuestas en ambos teoremas son las mismas. Notamos que las regiones están superpuestas. Por lo tanto, si comparamos la accesibilidad de los métodos simplificado de Newton (4.1) y de Steffensen, desde el punto de vista teórico de los dominios de parámetros asociados a los teoremas 4.3 y 4.1, concluimos lo mismo que antes: la accesibilidad del método de Steffensen se reduce considerablemente con respecto a la del método simplificado de Newton (4.1).

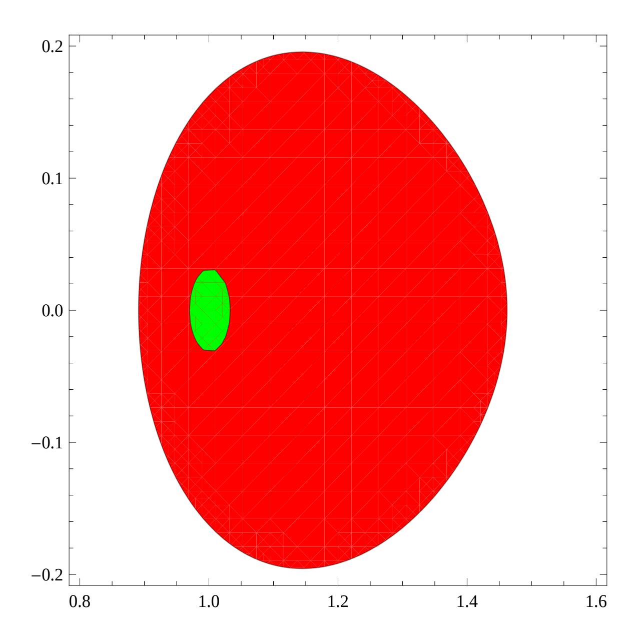


Figura 4.3: Regiones de accesibilidad de la solución $z = 1$ de la ecuación compleja $F(z) = z^3 - 1 = 0$ para los métodos simplificado de Newton (región roja) y de Steffensen (región verde), según los teoremas 4.3 y 4.1, respectivamente.

Parece entonces claro que podemos aprovechar la accesibilidad del método simplificado de Newton (4.1) para predecir puntos de salida para el método de Steffensen que garanticen la convergencia de éste al empezar en ellos.

4.3. Método iterativo híbrido (predictor-corrector)

A partir de lo visto anteriormente, ahora nos planteamos construir una modificación del método de Steffensen que mejore su accesibilidad, utilizando para esto el método simplificado de Newton (4.1).

4.3.1. Construcción del método

Como se observa en la figura 4.4, el dominio de parámetros del método de Steffensen está contenido en el dominio de parámetros del método simplificado de Newton (4.1). Por tanto, las condiciones exigidas al método de Steffensen para garantizar su convergencia semilocal son evidentemente más restrictivas que las exigidas al método simplificado de Newton (4.1).

Tratamos entonces de asegurar que, para un par inicial (δ_0, θ_0) que satisfaga la condición (4.8), es decir que esté dentro del dominio de parámetros del método simplificado de Newton, obtengamos un par (δ, θ) que satisfaga las dos condiciones dadas en (4.4), después de realizar un cierto número N_0 de iteraciones con el método simplificado de Newton, de manera que estemos en condiciones de poder garantizar la convergencia del método de Steffensen al empezar en la iteración N_0 que se obtiene mediante el método simplificado de Newton (4.1). De este modo, se puede considerar el par $(\delta_{N_0}, \theta_{N_0})$ asociado a la iteración N_0 obtenida con el método simplificado de Newton (4.1) como el par (δ, θ) asociado a la iteración inicial x_0 del método de Steffensen.

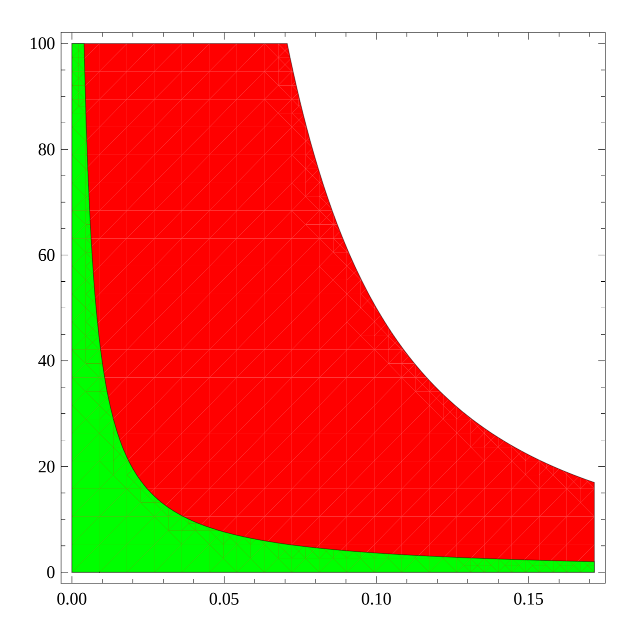


Figura 4.4: Dominios de parámetros de los métodos simplificado de Newton (región roja) y de Steffensen (región verde) asociados a los teoremas 4.3 y 4.1, respectivamente.

Nuestro objetivo principal es entonces construir una sencilla modificación del método de Steffensen de manera que este método converja al empezar en los mismos puntos de salida que el método simplificado de Newton (4.1). Consideramos entonces el método iterativo híbrido (predictor-corrector) dado por el siguiente algoritmo

$$\begin{cases} \begin{cases} \text{dado } z_0 \text{ en } \Omega, \\ z_{i+1} = z_i - [F'(z_0)]^{-1}F(z_i), \quad i = 0, 1, \ldots, N_0 - 1, \end{cases} \\ \begin{cases} x_0 = z_{N_0}, \\ x_{n+1} = x_n - [x_n, x_n + F(x_n); F]^{-1}F(x_n), \quad n \geq 0, \end{cases} \end{cases} \tag{4.9}$$

donde z_0 satisface (4.8), mientras que $x_0 = z_{N_0}$ satisface (4.4). Para que este algoritmo sea convergente nos planteamos dos cuestiones:

1. Localizar z_0 de manera que el método predictor, el método simplificado de Newton (4.1), sea convergente.

2. Utilizando la convergencia del método predictor, calcular un valor N_0 tal que z_{N_0} sea considerado como punto de salida del método corrector, el método de Steffensen, de manera que podamos asegurar la convergencia del método de Steffensen saliendo desde $x_0 = z_{N_0}$.

La idea es usar el método simplificado de Newton (4.1) durante un número finito de pasos N_0 hasta que $z_{N_0} = x_0$ cumpla las condiciones dadas en (4.4) y aplicar después el método de Steffensen en vez del método simplificado de Newton (4.1). La clave del problema reside entonces en garantizar la existencia de N_0.

4.3.2. Convergencia semilocal del método

A continuación, vamos a estudiar la convergencia semilocal del método híbrido (4.9). Necesitamos entonces conocer la evolución de los parámetros asociados a las aproximaciones que vamos obteniendo como iteraciones del método simplificado de Newton (4.1), para lo que nos apoyamos en el siguiente resultado.

Teorema 4.4. *En las condiciones del teorema 4.3, consideramos el polinomio (4.5) y sean t^* y t^{**} sus dos raíces reales positivas y tales que $t^* \leq t^{**}$. Entonces, para el método simplificado de Newton (4.1), se obtienen las siguientes cotas de error:*

(i) Si $t^* < t^{**}$,

$$\frac{((t^{**}-t^*)t^*)^{n+1}}{(t^{**})^{n+1}-(t^*)^{n+1}} < t^* - t_n < \frac{t^*(t^{**}-t^*)(t^*C\theta_0)^n}{t^{**}-t^*(t^*C\theta_0)^n}.$$

(ii) Si $t^* = t^{**}$,

$$\left(\frac{1}{2}\right)^n t^* \leq t^* - t_n \leq t^*.$$

Demostración. Demostramos en primer lugar el apartado (i). A partir de (4.5), como $t^* < t^{**}$, tenemos

$$p(t) = \frac{C}{2}(t^*-t)(t^{**}-t).$$

Si denotamos $a_n = t^* - t_n$ y $b_n = t^{**} - t_n$, para todo $n \geq 0$, entonces $p(t_n) = \frac{C}{2}a_n b_n$, $p'(t_0) = -\frac{1}{\theta_0}$ y

$$a_{n+1} = t^* - t_{n+1} = t^* - t_n - \theta_0 p(t_n) = a_n - \frac{C\theta_0}{2}a_n b_n,$$

$$b_{n+1} = t^{**} - t_{n+1} = t^{**} - t_n - \theta_0 p(t_n) = b_n - \frac{C\theta_0}{2}a_n b_n.$$

Como

$$\frac{a_{n+1}}{b_{n+1}} = \frac{a_n}{b_n}\frac{2-C\theta_0(t^{**}-t_n)}{2-C\theta_0(t^*-t_n)}$$

y la función

$$\psi(t) = \frac{2-C\theta_0(t^{**}-t)}{2-C\theta_0(t^*-t)}$$

es estrictamente creciente, entonces

$$\frac{t^*}{t^{**}} = \psi(0) = \psi(t_0) \leq \psi(t) \leq \psi(t^*) = C\theta_0 t^*,$$

de manera que

$$\frac{a_{n+1}}{b_{n+1}} = \frac{a_n}{b_n}\psi(t_n) \leq \frac{a_n}{b_n}\psi(t^*) \leq \frac{a_{n-1}}{b_{n-1}}\psi(t^*)^2 \leq \cdots \leq \frac{a_0}{b_0}\psi(t^*)^{n+1} = \frac{t^*}{t^{**}}(C\theta_0 t^*)^{n+1},$$

$$\frac{a_{n+1}}{b_{n+1}} = \frac{a_n}{b_n}\psi(t_n) \geq \frac{a_n}{b_n}\psi(t_0) \geq \frac{a_{n-1}}{b_{n-1}}\psi(t_0)^2 \geq \cdots \geq \frac{a_0}{b_0}\psi(t_0)^{n+1} = \left(\frac{t^*}{t^{**}}\right)^{n+2}.$$

Como $b_{n+1} = (t^{**} - t^*) + a_{n+1}$, obtenemos

$$\frac{(t^{**} - t^*)(t^*)^{n+2}}{(t^{**})^{n+2} - (t^*)^{n+2}} < t^* - t_{n+1} < \frac{t^*(t^{**} - t^*)(t^* C\theta_0)^{n+1}}{t^{**} - t^*(t^* C\theta_0)^{n+1}}.$$

En segundo lugar vemos (ii). Partiendo de (4.5), si $t^* = t^{**}$, entonces $a_n = b_n$ y $p(t_n) = \frac{C}{2} a_n^2$, de manera que

$$\frac{a_{n+1}}{a_n} = 1 - \frac{C\theta_0}{2} a_n.$$

Ahora, teniendo en cuenta que la función

$$\varphi(t) = 1 - \frac{C\theta_0}{2}(t^* - t)$$

es estrictamente creciente, se sigue que $\frac{1}{2} = \varphi(t_0) \leq \varphi(t) \leq \varphi(t^*) = 1$ y, como $a_{n+1} = \varphi(t_n) a_n$, entonces

$$a_{n+1} \leq a_n \leq \cdots \leq a_0 \leq t^* - t_0 = t^*,$$
$$a_{n+1} \geq \frac{1}{2} a_n \geq \cdots \geq \left(\frac{1}{2}\right)^{n+1} a_0 = \left(\frac{1}{2}\right)^{n+1} t^*.$$

En consecuencia,

$$\left(\frac{1}{2}\right)^{n+1} t^* \leq t^* - t_{n+1} \leq t^*. \quad \blacksquare$$

Notemos que a partir del resultado anterior resulta evidente que el método simplificado de Newton (4.1) tiene al menos convergencia lineal.

Método predictor: el método simplificado de Newton

Consideramos la situación inicial a partir del método simplificado de Newton (4.1). Dada la aproximación inicial z_0, partimos del par (δ_0, θ_0) definido a partir de (H1) y (H2):

$$\|\overline{\Gamma}_0\| = \|[F'(z_0)]^{-1}\| \leq \theta_0 \qquad \text{y} \qquad \|F(z_0)\| \leq \delta_0.$$

Para la convergencia del método debe verificarse la condición (4.8):

$$C\delta_0\theta_0^2 \leq \frac{1}{2}.$$

Iterando, se van definiendo los pares (δ_n, θ_n), a partir del cumplimiento de la condiciones (H1) y (H2), de modo que

$$\|F(z_n)\| \leq p(t_n) = \delta_n \qquad \Leftrightarrow \qquad \frac{C}{2} a_n b_n = \delta_n \qquad \Leftrightarrow \qquad \frac{C^2}{2} a_n b_n = C\delta_n,$$

donde $a_n = t^* - t_n$ y $b_n = t^{**} - t_n$, siendo t^* y t^{**} $(t^* \leq t^{**})$ las dos raíces positivas de (4.5). Por el teorema 4.4, tenemos

$$\begin{aligned} C\delta_n = Cp(t_n) &< \frac{C^2}{2} \frac{t^*\, t^{**}(t^{**} - t^*)}{t^{**} - t^*(t^* C\theta_0)^n} (t^* C\theta_0)^n \\ &= \frac{C t^* t^{**} \sqrt{1 - 2C\delta_0\theta_0^2}}{\theta_0 \left(t^{**} - t^*(t^* C\theta_0)^n\right)} (t^* C\theta_0)^n \\ &= \frac{C t^* t^{**} (t^* C\theta_0)^n}{\theta_0 (t^{**} - t^*(t^* C\theta_0)^n)} (1 - t^* C\theta_0). \end{aligned} \tag{4.10}$$

A continuación, expresando θ_n en función de θ_0 y considerando

$$\|I - \overline{\Gamma}_0 F'(z_n)\| \leq \|\overline{\Gamma}_0\| \|F'(z_0) - F'(z_n)\| \leq C\theta_0 \|z_0 - z_n\| \leq C\theta_0 t^*,$$

obtenemos

$$\|\overline{\Gamma}_n\| = \| \left[F'(z_n)\right]^{-1} \| \leq \theta_n, \quad \text{donde} \quad \theta_n = \frac{\theta_0}{1 - C\theta_0 t^*} = \theta.$$

Con lo que, tras aplicar este método, obtenemos el par $(\delta_n, \theta_n) = (\delta_n, \theta)$ para algún $n \in \mathbb{N}$.

Método corrector: el método de Steffensen

El par $(\delta_n, \theta_n) = (\delta_n, \theta)$ debe verificar ahora las condiciones de convergencia dadas en (4.4) para el método de Steffensen:

$$C\delta\theta \leq 2 \qquad \text{y} \qquad \ell\delta b^2 \leq \frac{1}{2},$$

donde $\delta = \delta_n$, $b = \frac{2\theta}{2 - C\delta\theta}$ y $\ell = C\left(1 + \frac{1}{b}\right)$. Por tanto, sustituyendo los parámetros δ y θ del método predictor en (4.4), tenemos:

- Si $C\delta\theta \leq 2$, entonces $C\delta_n \leq \frac{2}{\theta}$,
- Si $\ell\delta b^2 \leq \frac{1}{2}$, entonces $\frac{1}{2} \geq \ell\delta_n b^2 = C\left(1 + \frac{2 - C\delta_n\theta}{2\theta}\right)\delta_n\left(\frac{2\theta}{2 - C\delta_n\theta}\right)^2$, que es equivalente a

$$5\theta^2 (C\delta_n)^2 - 4\theta(3 + 2\theta)(C\delta_n) + 4 \geq 0$$

  y se verifica si

$$C\delta_n \leq P, \quad \text{donde} \quad P = \frac{2(3 + 2\theta)\theta - 4\sqrt{\theta^2 + 3\theta + 1}}{5\theta^2}.$$

Luego,

$$C\delta_n \leq \min\left\{\frac{2}{\theta}, P\right\} = \begin{cases} P & \text{si} \quad \theta \leq 2.5303\ldots \\ \dfrac{2}{\theta} & \text{si} \quad \theta \geq 2.5303\ldots \end{cases} \tag{4.11}$$

Definición de N_0

Ahora vamos a definir N_0 de manera que indique cuando se da el salto del método predictor al método corrector. Notemos que la sucesión $\{\delta_n\}$ es decreciente, con lo que buscamos el primer valor de n en (4.10) que verifique (4.11).

- Si $\theta \leq 2.5303\ldots$, entonces

$$C\delta_n < \frac{Ct^* t^{**} (C\theta_0 t^*)^n}{\theta_0 (t^{**} - t^* (t^* C\theta_0)^n)} (1 - C\theta_0 t^*) < P.$$

  Tomando logaritmos, encontramos un valor $N_0 \in \mathbb{N}$, de manera que el par $(\delta_{N_0}, \theta_{N_0})$ satisfaga las dos condiciones de convergencia dadas en (4.4) para el método corrector. Así, obtenemos

$$N_0 \geq \frac{\log\left(\dfrac{P\theta_0 t^{**}}{t^*(P\theta_0 + Ct^{**}(1 - t^* C\theta_0))}\right)}{\log(t^* C\theta_0)}.$$

Luego

$$N_0 = 1 + \left[\frac{\log\left(\frac{P\theta_0 t^{**}}{t^*(P\theta_0 + Ct^{**}(1 - C\theta_0 t^*))}\right)}{\log(C\theta_0 t^*)} \right].$$

- Si $\theta \geq 2.5303\ldots$, entonces

$$C\delta_n < \frac{Ct^* t^{**}(t^* C\theta_0)^n}{\theta_0(t^{**} - t^*(t^* C\theta_0)^n)}(1 - t^* C\theta_0) < \frac{2}{\theta}.$$

  Tomando de nuevo logaritmos y procediendo como en el caso anterior, encontramos un valor $N_0 \in \mathbb{N}$, de manera que el par $(\delta_{N_0}, \theta_{N_0})$ satisfaga las dos condiciones de convergencia dadas en (4.4) para el método corrector. Así, obtenemos

$$N_0 = 1 + \left[\frac{\log\left(\frac{2\theta_0 t^{**}}{t^*(2\theta_0 + C\theta t^{**}(1 - C\theta_0 t^*))}\right)}{\log(C\theta_0 t^*)} \right].$$

Por tanto, podemos asegurar que el método corrector es convergente partiendo del punto $x_0 = z_{N_0}$.

A partir de los comentarios y resultados previos, queda probado el siguiente teorema de convergencia semilocal para el método iterativo híbrido (predictor-corrector) dado en (4.9).

Teorema 4.5. *Sean X un espacio de Banach y $F : \Omega \subset X \to X$ un operador una vez continuamente diferenciable Fréchet definido en un conjunto abierto convexo no vacío Ω. Suponemos que se satisfacen las condiciones (H1)–(H3) y que $B(z_0, t^*) \subset \Omega$, donde t^* es la menor raíz positiva del polinomio (4.5). Si además $C\delta_0\theta_0^2 \leq \frac{1}{2}$, entonces existe $N_0 \in \mathbb{N}$ de forma que el método (4.9) es convergente con*

$$N_0 = \begin{cases} 1 + \left[\frac{\log\left(\frac{P\theta_0 t^{**}}{t^*(P\theta_0 + Ct^{**}(1 - C\theta_0 t^*))}\right)}{\log(C\theta_0 t^*)} \right] & si \quad \theta \leq 2.5303\ldots, \\ 1 + \left[\frac{\log\left(\frac{2\theta_0 t^{**}}{t^*(2\theta_0 + C\theta t^{**}(1 - C\theta_0 t^*))}\right)}{\log(C\theta_0 t^*)} \right] & si \quad \theta \geq 2.5303\ldots, \end{cases} \tag{4.12}$$

y $\theta = \dfrac{\theta_0}{1 - C\theta_0 t^*}$.

4.4. Aplicación

En principio, la aplicabilidad del método de Steffensen es menor que la del método simplificado de Newton (4.1). Veamos que esto es así en la siguiente aplicación, donde comprobamos

que no podemos aplicar el método de Steffensen para aproximar una solución del problema discreto, que surge de discretizar una ecuación integral no lineal de Hammerstein, porque no se cumplen las condiciones de convergencia del teorema 4.1. Sin embargo, apoyándonos en el método híbrido (4.9), vemos que podemos aplicar el método de Steffensen a partir de cierta iteración.

Consideramos la siguiente ecuación integral no lineal mixta de tipo Hammerstein

$$x(s) = 1 + \int_0^1 G(s,t)\, x(t)^2\, dt, \quad s \in [0,1], \tag{4.13}$$

donde $x \in C[0,1]$, $t \in [0,1]$ y el núcleo G es la función de Green en $[0,1] \times [0,1]$.

A continuación, usamos un proceso de discretización que transforma (4.13) en un problema finito dimensional, tal y como se hizo en la sección 1.6.1, de manera que obtenemos el siguiente sistema de ecuaciones no lineales:

$$F(\mathbf{x}) \equiv \mathbf{x} - \mathbf{1} - A\,\hat{\mathbf{x}} = 0, \qquad F : \mathbb{R}^8 \longrightarrow \mathbb{R}^8, \tag{4.14}$$

donde

$$\mathbf{x} = (x_1, x_2, \ldots, x_8)^T, \qquad \mathbf{1} = (1, 1, \ldots, 1)^T, \qquad A = (a_{ij})_{i,j=1}^8, \qquad \hat{\mathbf{x}} = (x_1^2, x_2^2, \ldots, x_8^2)^T.$$

Además , $F'(\mathbf{x}) = I - 2A\,\mathrm{diag}\{\mathbf{x}\}$.

Eligiendo como punto de salida $\mathbf{z_0} = (1.7, 1.7, \ldots, 1.7)^T$ y la norma del máximo, obtenemos $\delta_0 = 0.6713\ldots$, $\theta_0 = 1.6549\ldots$, $C = 0.2471\ldots$, $C\delta_0\theta_0^2 = 0.4543\ldots < \frac{1}{2}$. Por tanto, podemos aplicar el método simplificado de Newton (4.1) para resolver el sistema (4.14), puesto que se verifica la condición (4.8) del teorema 4.3. Por contra, no podemos aplicar el método de Steffensen, ya que no se verifica la segunda condición de (4.4) del teorema 4.1, puesto que

$$\ell\delta b^2 = 0.9286\ldots > \frac{1}{2},$$

donde $\delta = \delta_0$, $b = 1.9182\ldots$ y $\ell = 0.3759\ldots$

Como el método simplificado de Newton (4.1) es convergente por el teorema 4.3, lo aplicamos para obtener la aproximación numérica $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_8^*)^T$ de la solución del sistema (4.14) y mostrada en la tabla 4.1, después de 20 iteraciones y usando como criterio de parada $\|\mathbf{z_n} - \mathbf{z_{n-1}}\| < 10^{-16}$. En la tabla 4.2 se muestran los errores $\|\mathbf{z_n} - \mathbf{x}^*\|$ obtenidos con el mismo criterio de parada. Notemos que el vector dado en la tabla 4.1 es una buena aproximación de la solución del sistema (4.14), puesto que $\|F(\mathbf{x}^*)\| \leq$ constante $\times\, 10^{-16}$. Mostramos la sucesión $\{\|F(\mathbf{z_n})\|\}$ en la tabla 4.2.

i	x_i^*	i	x_i^*	i	x_i^*	i	x_i^*
1	1.012239 . . .	3	1.118079 . . .	5	1.159804 . . .	7	1.058428 . . .
2	1.058428 . . .	4	1.159804 . . .	6	1.118079 . . .	8	1.012239 . . .

Tabla 4.1: Aproximación de la solución $\mathbf{x}^*$ de (4.14)

Por otra parte, si aplicamos el método híbrido (4.9) para aproximar la solución dada en la tabla 4.1 con el mismo punto de salida, sólo necesitamos calcular el valor de N_0, mediante

n	$\|\mathbf{z_n} - \mathbf{x}^*\|$	$\|F(\mathbf{z_n})\|$
0	$6.8776\ldots \times 10^{-1}$	$6.7130\ldots \times 10^{-1}$
1	$6.1560\ldots \times 10^{-2}$	$4.6960\ldots \times 10^{-2}$
2	$1.1811\ldots \times 10^{-2}$	$8.9776\ldots \times 10^{-3}$
3	$2.0931\ldots \times 10^{-3}$	$1.5942\ldots \times 10^{-3}$
4	$3.7465\ldots \times 10^{-4}$	$2.8588\ldots \times 10^{-4}$
5	$6.6908\ldots \times 10^{-5}$	$5.0952\ldots \times 10^{-5}$
$\vdots$	$\vdots$	$\vdots$
19	$2.2532\ldots \times 10^{-15}$	$1.7180\ldots \times 10^{-15}$

Tabla 4.2: Errores absolutos obtenidos con el método simplificado de Newton y $\{\|F(\mathbf{z_n})\|\}$

el teorema 4.5, teniendo en cuenta el valor $\theta = 5.4777\ldots$ Si nos fijamos en la fórmula (4.12), obtenemos $N_0 = 1$. Por tanto, después de una iteración del método simplificado de Newton (4.1), podemos aplicar el método de Steffensen, obteniendo la aproximación numérica de la solución dada en la tabla 4.1 tras realizar cuatro iteraciones más con este último método. En la tabla 4.3 se muestran los errores $\|\mathbf{x_n} - \mathbf{x}^*\|$ cuando usamos el criterio de parada $\|\mathbf{x_n} - \mathbf{x_{n-1}}\| < 10^{-16}$, así como la sucesión $\{\|F(\mathbf{x_n})\|\}$.

n	$\|\mathbf{x_n} - \mathbf{x}^*\|$	$\|F(\mathbf{x_n})\|$
0	$6.8776\ldots \times 10^{-1}$	$6.7130\ldots \times 10^{-1}$
1	$6.1560\ldots \times 10^{-2}$	$4.6960\ldots \times 10^{-2}$
2	$8.2751\ldots \times 10^{-4}$	$6.3230\ldots \times 10^{-4}$
3	$1.4582\ldots \times 10^{-7}$	$1.1158\ldots \times 10^{-7}$
4	$4.4982\ldots \times 10^{-15}$	$3.4428\ldots \times 10^{-15}$

Tabla 4.3: Errores absolutos obtenidos con el método híbrido (4.9) y $\{\|F(\mathbf{x_n})\|\}$

Capítulo 5

Situación (no)-diferenciable

El estudio realizado en el capítulo anterior se restringe a la resolución de ecuaciones $F(x) = 0$ en las que el operador F es diferenciable Fréchet. Evidentemente, una característica que tiene el método de Steffensen es que no necesita que el operador F sea diferenciable Fréchet, simplemente necesita que exista una diferencia dividida definida en el dominio Ω.

Habitualmente, para probar la convergencia semilocal del método de Steffensen, se imponen las siguientes dos condiciones básicas ([1],[2],[5]):

(I) Dados dos puntos cualesquiera $x, y \in \Omega$, distintos, existe la diferencia dividida de primer orden $[x, y; F]$.

(II) La diferencia dividida verifica la condición

$$\|[x, y; F] - [u, v; F]\| \leq K(\|x - u\| + \|y - v\|), \tag{5.1}$$

con $K \geq 0$, $x, y, u, v \in \Omega$, $x \neq y$ y $u \neq v$.

La condición (5.1) se puede suavizar considerando la condición

$$\|[x, y; F] - [u, v; F]\| \leq K(\|x - u\|^p + \|y - v\|^p), \tag{5.2}$$

con $K \geq 0$, $p \in (0, 1]$, $x, y, u, v \in \Omega$, $x \neq y$ y $u \neq v$. Las condiciones (5.1) y (5.2) dicen que la diferencia dividida de primer orden es, respectivamente, Lipschitz continua en Ω y (K, p)-Hölder continua en Ω.

Si ahora consideramos una diferencia dividida de primer orden para la que existe una función $\omega : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ no decreciente en sus dos argumentos y tal que

$$\|[x, y; F] - [u, v; F]\| \leq \omega(\|x - u\|, \|y - v\|), \tag{5.3}$$

con $x, y, u, v \in \Omega$, $x \neq y$ y $u \neq v$, es claro que obtenemos como casos particulares las condiciones (5.1) y (5.2) si $\omega(s, t) = K(s + t)$ o $\omega(s, t) = K(s^p + t^p)$, respectivamente. En estos casos, como $\omega(0, 0) = 0$, F es diferenciable (véase [37]). Así, los resultados de convergencia semilocal para el método de Steffensen dados habitualmente exigen veladamente que el operador F sea diferenciable.

En este capítulo nos planteamos la obtención de un resultado de convergencia semilocal para el método de Steffensen cuando se aplica a ecuaciones en las que el operador F puede ser no diferenciable. Para ello, consideramos las condiciones (I) y (5.3) teniendo en cuenta que $\omega(0, 0) \neq 0$ si el operador F es no diferenciable.

Por otra parte, dada la generalidad de las condiciones de convergencia que vamos a considerar, veremos, a partir del dominio de parámetros asociado al resultado de convergencia semilocal, que la accesibilidad del método de Steffensen va a ser restrictiva en determinadas condiciones. Para solventar esta dificultad, construiremos un método iterativo híbrido (predictor-corrector).

En la sección 5.1 obtenemos un resultado de convergencia semilocal para el método de Steffensen que permite aplicar este método a la resolución de ecuaciones con determinados operadores no diferenciables. Para ello, utilizamos la técnica de demostración de la convergencia semilocal ya vista en el capítulo 4, basada en relaciones de recurrencia, que proporciona novedosos resultados acerca de la convergencia semilocal. Después, analizamos el dominio de parámetros asociado al resultado obtenido, observando que es mejorable en determinadas situaciones. En la sección 5.2, pensando en la construcción de un método iterativo híbrido (predictor-corrector) que mejore la accesibilidad del método de Steffensen y que no utilice derivadas en su algoritmo, obtenemos un nuevo resultado de convergencia semilocal en las mismas condiciones, (I) y (5.3), para el método simplificado de Steffensen, cuyo algoritmo es

$$\begin{cases} \text{dado } z_0 \text{ en } \Omega, \\ z_{n+1} = z_n - [z_0, z_0 + F(z_0); F]^{-1} F(z_n), \quad n \geq 0, \end{cases} \tag{5.4}$$

y que utilizamos como método predictor en el método híbrido. Vemos que el dominio de parámetros de este método es menos restrictivo que el del método de Steffensen. Así, en la sección 5.3, construimos un método iterativo híbrido (predictor-corrector) que se beneficia del buen dominio de parámetros del método simplificado de Steffensen, el método predictor, y de la velocidad de convergencia del método de Steffensen, el método corrector. Terminamos con la sección 5.4, donde vemos dos aplicaciones en las que se aproximan las soluciones de dos sistemas no lineales, uno diferenciable y otro no diferenciable, mediante el método híbrido previamente construido, pero que no se pueden aproximar mediante el método de Steffensen.

5.1. Método corrector: el método de Steffensen

5.1.1. Convergencia semilocal

Tal y como hemos dicho en el capítulo anterior, resulta evidente que, en el estudio de la convergencia de la sucesión $\{x_n\}$ definida por el método de Steffensen, existirán todas las diferencias divididas de primer orden $[x_k, y_k; F]$, salvo que $x_k = y_k = x_k + F(x_k)$, en cuyo caso resulta evidente que x_k es una solución de la ecuación $F(x) = 0$ y, en este caso, tenemos $x_n = x_k$ para todo $n \geq k$, luego la sucesión $\{x_n\}$ es convergente a $x_k \equiv x^*$ solución de $F(x) = 0$.

Comenzamos con un lema técnico que utilizamos posteriormente.

Lema 5.1. *Sea $\{x_n\}$ la sucesión dada por el método de Steffensen. Si $x_{m-1} \neq x_m$ con $x_{m-1}, x_m \in \Omega$, entonces*

$$F(x_m) = ([x_m, x_{m-1}; F] - A_{m-1})(x_m - x_{m-1}), \text{ donde } A_{m-1} = [x_{m-1}, x_{m-1} + F(x_{m-1}); F].$$

DEMOSTRACIÓN. A partir de la sucesión dada por el método de Steffensen se sigue

$$F(x_{m-1}) + [x_{m-1}, x_{m-1} + F(x_{m-1}); F](x_m - x_{m-1}) = 0,$$

de manera que

$$\begin{aligned} F(x_m) &= F(x_m) - F(x_{m-1}) - [x_{m-1}, x_{m-1} + F(x_{m-1}); F](x_m - x_{m-1}) \\ &= [x_m, x_{m-1}; F](x_m - x_{m-1}) - [x_{m-1}, x_{m-1} + F(x_{m-1}); F](x_m - x_{m-1}) \\ &= ([x_m, x_{m-1}; F] - A_{m-1})\,(x_m - x_{m-1}). \quad \blacksquare \end{aligned}$$

A continuación, presentamos un resultado de convergencia semilocal para el método de Steffensen. Para ello, dados $x_0, x_0 + F(x_0) \in \Omega$, notemos que $x_0 \neq x_0 + F(x_0)$, ya que en otro caso x_0 es una solución x^* de $F(x) = 0$ y $x_n = x^*$, para todo $n \in \mathbb{N}$. Suponemos las siguientes condiciones:

- (C_1) $\|F(x_0)\| \leq \delta$,
- (C_2) existe $A_0^{-1} = [x_0, x_0 + F(x_0); F]^{-1} \in \mathcal{L}(X, X)$, para $x_0 \in \Omega$, y es tal que $\|A_0^{-1}\| \leq \beta$,
- (C_3) $\|[x, y; F] - [u, v; F]\| \leq \omega(\|x - u\|, \|y - v\|)$; $x, y, u, v \in D$; $x \neq y; u \neq v$, donde $\omega : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ es una función continua no decreciente en los dos argumentos.

Teorema 5.2. *Sean X un espacio de Banach y $F : \Omega \subset X \to X$ un operador no lineal definido en un dominio abierto convexo no vacío Ω. Suponemos que se cumplen las condiciones (C_1)–(C_3). Si la ecuación*

$$t = \frac{\beta\delta(1 - \beta\omega(t, t + \delta))}{1 - \beta\omega(t, t + \delta) - M} + M\delta, \tag{5.5}$$

donde $M = \beta\omega(\beta\delta, \delta)$, tiene al menos una raíz real positiva, y denotamos por R la raíz positiva más pequeña de (5.5),

$$M + \beta\omega(R, R + \delta) < 1 \tag{5.6}$$

y $B(x_0, R) \subset \Omega$, entonces el método de Steffensen, empezando en x_0, está bien definido y converge a una solución x^* de $F(x) = 0$. Además, la solución x^* y las iteraciones x_n pertenecen a $\overline{B(x_0, R)}$ y x^* es única en $\overline{B(x_0, R)} \cap \Omega$.

Demostración. Comenzamos probando que la sucesión $\{x_n\}$ está bien definida, es decir, $x_n \in B(x_0, R) \subset \Omega$, para todo $n \in \mathbb{N}$. Notemos que

$$R = \frac{\beta\delta}{1 - P} + M\delta, \tag{5.7}$$

donde $P = \dfrac{M}{1 - \beta\omega(R, R + \delta)} < 1$.

A partir de (C_1)–(C_2) se sigue que x_1 está bien definida y, por (5.7), tenemos

$$\|x_1 - x_0\| \leq \|A_0^{-1}\|\,\|F(x_0)\| \leq \beta\delta < R.$$

Luego, $x_1 \in B(x_0, R)$. Además, por el lema 5.1, $F(x_1) = ([x_1, x_0; F] - A_0)\,(x_1 - x_0)$ y, en consecuencia,

$$\begin{aligned} \|F(x_1)\| &\leq \|[x_1, x_0; F] - A_0\|\|x_1 - x_0\| \\ &\leq \omega(\|x_1 - x_0\|, \|F(x_0)\|)\|x_1 - x_0\| \\ &\leq \omega(\beta\delta, \delta)\|x_1 - x_0\| \\ &\leq M\delta \\ &< \delta. \end{aligned}$$

De nuevo, por (5.7), se sigue $\|x_1 + F(x_1) - x_0\| \leq \|x_1 - x_0\| + \|F(x_1)\| \leq \beta\delta + M\delta < R$ y $x_1 + F(x_1) \in B(x_0, R)$.

A continuación, teniendo en cuenta

$$\begin{aligned} \|I - A_0^{-1}A_1\| &\leq \|A_0^{-1}\|\,\|A_0 - A_1\| \\ &\leq \beta\omega(\|x_1 - x_0\|, \|x_1 + F(x_1) - x_0 - F(x_0)\|) \\ &\leq \beta\omega(\|x_1 - x_0\|, \|x_1 + F(x_1) - x_0\| + \|F(x_0)\|) \\ &\leq \beta\omega(\beta\delta, \beta\delta + \delta) \\ &\leq \beta\omega(R, R + \delta) \\ &< 1, \end{aligned}$$

vemos, por el lema de Banach (lema 1.22), que existe el operador A_1^{-1} y es tal que

$$\|A_1^{-1}\| \leq \frac{\beta}{1 - \beta\omega(R, R + \delta)}.$$

En consecuencia, como $P < 1$, se sigue

$$\|x_2 - x_1\| \leq \|A_1^{-1}\|\|F(x_1)\| \leq P\|x_1 - x_0\| < \beta\delta,$$

$$\|x_2 - x_0\| \leq \|x_2 - x_1\| + \|x_1 - x_0\| \leq (1 + P)\|x_1 - x_0\| = \frac{1 - P^2}{1 - P}\|x_1 - x_0\| < \frac{\beta\delta}{1 - P} < R.$$

Por lo tanto, $x_2 \in B(x_0, R)$. También, como

$$\begin{aligned} \|F(x_2)\| &\leq \|[x_2, x_1; F] - A_1\|\|x_2 - x_1\| \\ &\leq \omega(\|x_2 - x_1\|, \|F(x_1)\|)\|x_2 - x_1\| \\ &\leq \omega(\beta\delta, \delta)\|x_2 - x_1\| \\ &\leq M\delta \\ &< \delta, \end{aligned}$$

se sigue

$$\|x_2 + F(x_2) - x_0\| \leq \|x_2 - x_0\| + \|F(x_2)\| < \frac{\beta\delta}{1 - P} + M\delta = R$$

y $x_2 + F(x_2) \in B(x_0, R)$.

Ahora podemos demostrar por inducción matemática sobre n que se cumplen los siguientes tres ítems para $n \in \mathbb{N}$:

· El operador A_n^{-1} existe y es tal que $\|A_n^{-1}\| \leq \dfrac{\beta}{1 - \beta\omega(R, R + \delta)}$,

· $\|F(x_n)\| \leq \omega(\beta\delta, \delta)\|x_n - x_{n-1}\|$,

· $\|x_{n+1} - x_n\| \leq P\|x_n - x_{n-1}\| \leq P^n\|x_1 - x_0\| < \beta\delta$,

siempre que $A_i = [x_i, x_i + F(x_i); F]$ sea invertible y $x_{i+1}, x_{i+1} + F(x_{i+1}) \in B(x_0, R)$, para todo $i = 1, 2, \ldots, n-1$.

En primer lugar, por hipótesis, vemos

$$\begin{aligned}
\|I - A_0^{-1}A_n\| &\leq \|A_0^{-1}\|\|A_0 - A_n\| \\
&\leq \beta\omega(\|x_n - x_0\|, \|x_n - x_0 + F(x_n) - F(x_0)\|) \\
&\leq \beta\omega(\|x_n - x_0\|, \|x_n + F(x_n) - x_0\| + \|F(x_0)\|) \\
&\leq \beta\omega(R, R+\delta) \\
&< 1
\end{aligned}$$

y

$$\|A_n^{-1}\| \leq \frac{\beta}{1 - \beta\omega(R, R+\delta)}.$$

En segundo lugar, por el lema 5.1, $F(x_n) = ([x_n, x_{n-1}; F] - A_{n-1})(x_n - x_{n-1})$ y

$$\begin{aligned}
\|F(x_n)\| &\leq \|[x_n, x_{n-1}; F] - A_{n-1}\|\|x_n - x_{n-1}\| \\
&\leq \omega(\|x_n - x_{n-1}\|, \|F(x_{n-1})\|)\|x_n - x_{n-1}\| \\
&\leq \omega(\beta\delta, \delta)\|x_n - x_{n-1}\| \\
&\leq M\delta \\
&< \delta.
\end{aligned}$$

Como ya hemos indicado anteriormente, existen todas las diferencias divididas de primer orden utilizadas porque, en otro caso, obtendríamos $x_n = x_{n-1}$ o $F(x_n) = 0$, lo que indicaría que ya habríamos alcanzado una solución de $F(x) = 0$ y el resultado quedaría probado.

En tercer lugar, vemos que

$$\begin{aligned}
\|x_n - x_{n-1}\| &\leq \|A_n^{-1}\|\|F(x_n)\| \\
&\leq \frac{\beta\omega(\beta\delta, \delta)}{1 - \beta\omega(R, R+\delta)}\|x_n - x_{n-1}\| \\
&= P\,\|x_n - x_{n-1}\| \\
&\leq P^n\,\|x_1 - x_0\| \\
&< \beta\delta.
\end{aligned}$$

En consecuencia,

$$\begin{aligned}
\|x_{n+1} - x_0\| &\leq \sum_{i=0}^{n+1}\|x_i - x_{i-1}\| \\
&\leq (P^n + P^{n-1} + \cdots + P + 1)\|x_1 - x_0\| \\
&\leq \frac{1 - P^{n+1}}{1 - P}\|x_1 - x_0\| \\
&< \frac{\beta\delta}{1 - P} \\
&< R,
\end{aligned}$$

$$\|x_{n+1} + F(x_{n+1}) - x_0\| \leq \|x_{n+1} - x_0\| + \|F(x_{n+1})\| < \frac{\beta\delta}{1-P} + M\delta = R,$$

por (5.7) y ser $P < 1$. Luego, $x_{n+1}, x_{n+1} + F(x_{n+1}) \in B(x_0, R)$, para todo $n \in \mathbb{N}$.

Una vez probado que la sucesión $\{x_n\}$ está bien definida, vemos que es una sucesión de Cauchy. En efecto, como

$$\begin{aligned}
\|x_{n+j} - x_n\| &\leq \|x_{n+j} - x_{n+j-1}\| + \|x_{n+j-1} - x_{n+j-2}\| + \cdots + \|x_{n+1} - x_n\| \\
&\leq (P^{j-1} + P^{j-2} + \cdots + P + 1)\|x_{n+1} - x_n\| \\
&= \frac{1-P^j}{1-P}\,\|x_{n+1} - x_n\| \\
&< \frac{P^n}{1-P}\,\|x_1 - x_0\|,
\end{aligned}$$

para $j \geq 1$, y $P < 1$, es claro que $\{x_n\}$ es una sucesión de Cauchy. Por tanto, la sucesión $\{x_n\}$ es convergente. Ahora, si $\lim_n x_n = x^*$, vemos que x^* es una solución de $F(x) = 0$. Como

$$\|F(x_n)\| \leq \omega(\beta\delta, \delta)\,\|x_n - x_{n-1}\|$$

y $\|x_n - x_{n-1}\| \to 0$, cuando $n \to \infty$, se sigue fácilmente, por la continuidad de F, que $F(x^*) = 0$.

Finalmente, probamos la unicidad de la solución x^* en $\overline{B(x_0, R)}$. Supongamos entonces que tenemos otra solución $y^* \in \overline{B(x_0, R)}$, $y^* \neq x^*$, de la ecuación $F(x) = 0$. Sea $J = [y^*, x^*; F]$. Si J es inversible, entonces $x^* = y^*$, puesto que $J(y^* - x^*) = F(y^*) - F(x^*)$. Para ver que J es invertible, basta, por el lema de Banach (lema 1.22), con ver que $\|I - A_0^{-1}J\| < 1$. En efecto, si $x_0 \neq x_0 + F(x_0)$, por hipótesis tenemos

$$\begin{aligned}
\|I - A_0^{-1}J\| &\leq \|A_0^{-1}\|\|A_0 - J\| \\
&\leq \|A_0^{-1}\|\|[x_0, x_0 + F(x_0); F] - [y^*, x^*; F]\| \\
&\leq \beta\omega(\|y^* - x_0\|, \|x^* - x_0 - F(x_0)\|) \\
&\leq \beta\omega(\|y^* - x_0\|, \|x^* - x_0\| + \|F(x_0)\|) \\
&\leq \beta\omega(R, R + \delta) \\
&< 1.
\end{aligned}$$

Luego, el operador J^{-1} existe. ■

5.1.2. Accesibilidad

Una vez probada la convergencia semilocal del método de Steffensen, nuestro siguiente objetivo es ver cuál es el dominio de parámetros de este método.

Es conocido que el operador F es diferenciable si la función ω que aparece en (C_3) cumple $\omega(0,0) = 0$, [37]. En otro caso, el operador F puede ser no diferenciable. A continuación, vamos a analizar dos casos particulares de la función ω: uno diferenciable y otro no diferenciable.

Como hemos visto en la discretización de ecuaciones diferenciales e integrales, es conocido que la función ω que aparece en (C_3) es frecuentemente de la forma

$$\omega(s,t) = L + K(s+t), \quad L, K \geq 0. \tag{5.8}$$

En este caso, la ecuación (5.5) del teorema 5.2 se puede transformar en la siguiente ecuación cuadrática:

$$2K\beta t^2 + (M(1-2K\beta\delta) + \beta(L + K\delta(1-2\beta)) - 1)t - \delta(M^2 + (L\beta + K\beta\delta - 1)(M+\beta)) = 0, \tag{5.9}$$

donde $M = \beta(L + K\delta(1+\beta))$. Y la condición (5.6) se transforma en la condición

$$M + \beta(L + K(2R+\delta)) < 1, \tag{5.10}$$

donde R es la raíz positiva más pequeña de (5.9) siempre que exista.

Analizamos a continuación la ecuación (5.9) viendo cuándo tiene dos raíces reales positivas. La ecuación anterior tendrá dos raíces reales positivas si

$$M(1-2K\delta\beta) + \beta(L + K\delta(1-2\beta)) - 1 < 0, \tag{5.11}$$

$$\delta\left(M^2 + (L\beta + K\delta\beta - 1)(M+\beta)\right) < 0, \tag{5.12}$$

$$\begin{aligned}
\Delta &= (M(1-2K\delta\beta) + \beta(L + K\delta(1-2\beta)) - 1)^2 + 8K\delta\beta\,(M^2 + (L\beta + K\delta\beta - 1)(M+\beta)) \\
&= \left(M(1-2K\delta\beta) + \beta(L + K\delta(1-2\beta)) - 1 + \sqrt{-8K\delta\beta\,(M^2 + (L\beta + K\delta\beta - 1)(M+\beta))}\right) \\
&\quad \times\left(M(1-2K\delta\beta) + \beta(L + K\delta(1-2\beta)) - 1 - \sqrt{-8K\delta\beta\,(M^2 + (L\beta + K\delta\beta - 1)(M+\beta))}\right) \\
&> 0.
\end{aligned}$$

Observamos que los dos factores de Δ son < 0 si

$$M(1-2K\beta\delta) + \beta(L + K\delta(1-2\beta)) + \sqrt{-8K\beta\delta(M^2 + (L\beta + K\beta\delta - 1)(M+\beta))} < 1. \tag{5.13}$$

Además, se cumple (5.11) si se satisface (5.13). Por lo tanto, la ecuación (5.9) tendrá dos raíces reales positivas si se cumplen (5.12) y (5.13). Si la ecuación (5.9) tiene dos raíces reales positivas, entonces la raíz positiva más pequeña es:

$$R = \frac{1}{4K\beta}\left(1 - M(1-2K\beta\delta) - \beta(L + K\delta(1-2\beta)) - \sqrt{\Delta}\right). \tag{5.14}$$

Notemos que también podemos considerar que la ecuación (5.9) tenga una raíz doble sin más que tener en cuenta la desigualdad no estricta en (5.13).

A continuación, sustituimos el valor de R en (5.10) y vemos que (5.10) se cumple si

$$1 - L\beta - (1 + 2K\delta\beta)M - K\delta\beta(1+2\beta) + \sqrt{\Delta} > 0. \tag{5.15}$$

Destacamos que no consideramos la posibilidad de que la ecuación cuadrática (5.9) tenga una raíz real positiva y otra negativa porque, en este caso, la raíz real positiva nunca cumple la condición (5.15).

Teniendo en cuenta lo anterior, enunciamos el siguiente resultado, cuya demostración se sigue fácilmente sin más que cumplir las hipótesis del teorema 5.2 cuando $\omega(s,t)$ es la función definida en (5.8).

Corolario 5.3. *Sean X un espacio de Banach y $F:\Omega\subset X\to X$ un operador no lineal definido en un dominio abierto convexo no vacío Ω. Supongamos que se cumplen las condiciones (C_1)–(C_3), siendo ω la función que se define en (5.8). Además, si se cumplen (5.12), (5.13), (5.15) y $B(x_0,R)\subset\Omega$, donde R está definido en (5.14), entonces el método de Steffensen, empezando en x_0, está bien definido y converge a una solución x^* de $F(x)=0$. Además, la solución x^* y las iteraciones x_n pertenecen a $\overline{B(x_0,R)}$ y x^* es única en $\overline{B(x_0,R)}\cap\Omega$.*

Caso diferenciable

A continuación, analizamos la accesibilidad del método de Steffensen desde el punto de vista teórico de los dominios de parámetros asociados al resultado anterior de convergencia semilocal. Para ello, distinguimos dos casos: el caso diferenciable, $L=0$, y el caso no diferenciable, $L\neq 0$.

En la figura 5.1 podemos visualizar el dominio de parámetros asociado al corolario 5.3 cuando F es diferenciable (caso $L=0$). Así, el dominio de parámetros es la región del plano cuyos puntos representan los parámetros correspondientes a los puntos de salida a partir de los cuales está garantizada la convergencia del método de Steffensen, indicando entonces a partir de qué puntos de salida está garantizada la convergencia del método bajo las condiciones (C_1)–(C_3), donde ω es la función definida en (5.8) con $L=0$. Para representarlo gráficamente, consideramos el plano xy, con $x=\beta$ (eje de abscisas) e $y=K\delta$ (eje de ordenadas), y coloreamos los valores de los parámetros que verifican las condiciones (5.12), (5.13) y (5.15), cuando $L=0$, y que se imponen en el corolario 5.3. Observamos que las condiciones iniciales (C_1) y (C_2), exigidas al punto de salida x_0, definen los parámetros δ y β, mientras que la condición (C_3), exigida al operador F, define el parámetro fijo K.

Por otra parte, consideramos el análisis de la convergencia semilocal del método de Steffensen mediante el principio de la mayorante de Kantorovich, sección 4.1.1, y representamos gráficamente el dominio de parámetros asociado al teorema 4.1, de manera que podamos compararlo con el asociado al corolario 5.3 cuando $L=0$ y ver cuál es mayor. Para ello, tenemos que representar los mismos valores en los ejes x e y del plano en el que vamos a representar los dos dominios de parámetros. En consecuencia, tenemos que escribir θ y C en función de β y K, de manera que así representamos los valores de los inversos de las diferencias divididas en el eje x y, en el eje y, el producto de δ por la constante de Lipschitz K para la diferencia dividida. Procedemos entonces como se detalla a continuación.

Por una parte, como

$$\|I-A_0^{-1}F'(x_0)\|\leq\|A_0^{-1}\|\|A_0-F'(x_0)\|\leq\beta\|A_0-[x_0,x_0;F]\|\leq K\beta\|F(x_0)\|=K\beta\delta,$$

se sigue, por el lema de Banach (lema 1.22), que si $K\beta\delta<1$, existe $[F'(x_0)]^{-1}$ y es tal que

$$\|[F'(x_0)]^{-1}\|\leq\frac{\beta}{1-K\beta\delta}.$$

Así, $\theta=\dfrac{\beta}{1-K\beta\delta}$.

Por otra parte, como

$$\|F'(x)-F'(y)\|=\|[x,x;F]-[y,y;F]\|\leq 2K\|x-y\|,$$

se sigue que $C=2K$.

En consecuencia, para que se cumplan las dos condiciones dadas en (4.4), teorema 4.1, se tienen que cumplir las siguientes tres condiciones

$$K\beta\delta < 1, \qquad \frac{2K\beta\delta}{1-K\beta\delta} \leq 2 \qquad \text{y} \qquad 2K\beta\delta\left(\frac{1-\beta(2K\delta-1)}{(1-K\beta\delta)^2}\right) \leq \frac{1}{2},$$

que se reducen a

$$2K\beta\delta \leq 1 \qquad \text{y} \qquad 4(K\beta\delta)^2 - 4K\beta\delta(2+\beta(1-2K\delta)) + 1 \geq 0.$$

Y, por tanto, ya estamos en condiciones de poder comparar los dominios de parámetros asociados al corolario 5.3 cuando $L = 0$ y al teorema 4.1.

En la figura 5.1 vemos claramente que el dominio de parámetros asociado al corolario 5.3 cuando $L = 0$ es mayor que el asociado al teorema 4.1. Notamos que las regiones están superpuestas. Por lo tanto, utilizando la técnica de demostración de la convergencia semilocal desarrollada en este capítulo, mejoramos el dominio de parámetros que se obtiene mediante la técnica clásica del principio de la mayorante de Kantorovich desarrollada en el capítulo anterior.

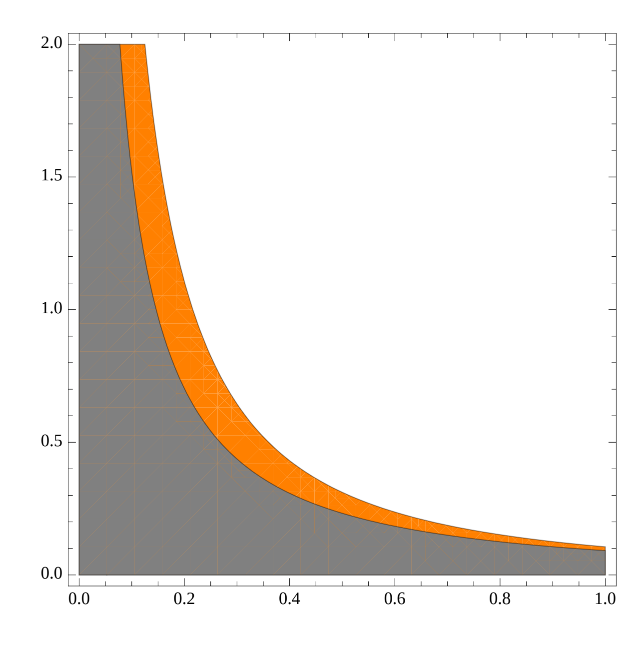


Figura 5.1: Dominios de parámetros del método de Steffensen asociados al corolario 5.3 cuando $L = 0$ (anaranjado) y al teorema 4.1 (gris).

Caso no diferenciable

Si ahora consideramos el corolario 5.3 con $L \neq 0$, al fijarnos en las condiciones (5.12), (5.13), (5.15), vemos que el valor de L está libre. Al visualizar entonces en la figura 5.2 el dominio de parámetros asociado a dicho corolario, con $L \neq 0$, podemos decir que el dominio de parámetros es mayor cuanto menor es el valor de L, obteniéndose como situación óptima el caso diferenciable ($L = 0$). Notamos que las regiones están superpuestas.

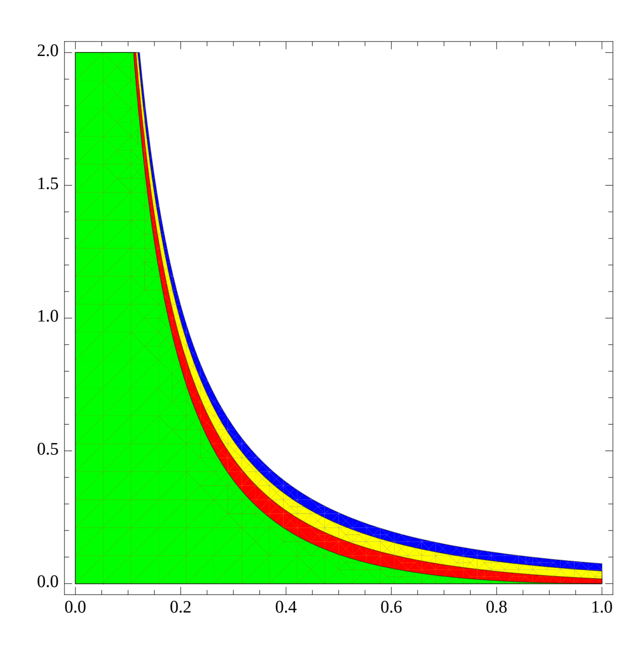


Figura 5.2: Dominios de parámetros del método de Steffensen asociados al corolario 5.3 cuando $L = \frac{1}{2}, \frac{1}{3}, \frac{1}{5}, \frac{1}{10}$ (verde, rojo, amarillo y azul, respectivamente).

5.1.3. Aplicación

A continuación ilustramos el estudio realizado anteriormente con dos sistemas de ecuaciones no lineales, uno diferenciable y otro no diferenciable, que surgen de la discretización de ecuaciones integrales de Hammerstein de la forma (1.41), presentadas en la sección 1.6.1. Para el sistema diferenciable, veremos que podemos garantizar la convergencia semilocal del método de Steffensen mediante el teorema 5.2 de este capítulo, pero no mediante el terorema 4.1 del capítulo anterior. Para el sistema no diferenciable, veremos que podemos garantizar la convergencia semilocal del método de Steffensen mediante el teorema 5.2.

Así, consideramos (1.41) con

$$H(t, x(t)) = \delta x(t)^2 + \mu|x(t)|, \quad \delta, \mu \in \mathbb{R},$$

y la trasformamos, mediante un proceso de discretización, tal y como se hace en la sección 1.6.1, en el siguiente sistema de ecuaciones no lineales:

$$F(\mathbf{x}) \equiv \mathbf{x} - \mathbf{f} - A\left(\delta\,\hat{\mathbf{x}} + \mu\,\tilde{\mathbf{x}}\right) = 0, \qquad F : \mathbb{R}^8 \longrightarrow \mathbb{R}^8, \tag{5.16}$$

donde $\hat{\mathbf{x}} = (x_1^2, x_2^2, \ldots, x_8^2)^T$, $\tilde{\mathbf{x}} = (|x_1|, |x_2|, \ldots, |x_8|)^T$ y $\delta, \mu \in \mathbb{R}$. Además,

$$[\mathbf{u}, \mathbf{v}; F] = I - A\left(\delta \mathrm{diag}\{\mathbf{z}\} + \mu \mathrm{diag}\{\mathbf{w}\}\right),$$

donde $\mathbf{z} = (z_1, z_2, \ldots, z_8)^T$ con $z_i = u_i + v_i$, para todo $i = 1, 2, \ldots, 8$, y $\mathbf{w} = (w_1, w_2, \ldots, w_8)^T$ con $w_i = \frac{|u_i|-|v_i|}{u_i-v_i}$, para todo $i = 1, 2, \ldots, 8$, de manera que $L = 2|\mu|\|A\|$ y $K = |\delta|\|A\|$.

Sistema de ecuaciones no lineales diferenciable

Si $\mathbf{f} = \mathbf{2} = (2, 2, \ldots, 2)^T$, $\delta = \frac{3}{4}$ y $\mu = 0$, entonces el sistema (5.16) se reduce a

$$F(\mathbf{x}) \equiv \mathbf{x} - \mathbf{2} - \frac{3}{4} A\,\hat{\mathbf{x}} = 0, \qquad F : \mathbb{R}^8 \longrightarrow \mathbb{R}^8. \tag{5.17}$$

En este caso, $[\mathbf{u}, \mathbf{v}; F] = I - \frac{3}{4}A\,\mathrm{diag}\{\mathbf{z}\}$ y el sistema de ecuaciones no lineales es diferenciable ($\mu = 0$).

Eligiendo $\mathbf{x_0} = \left(\frac{7}{5}, \frac{7}{5}, \ldots, \frac{7}{5}\right)^T$ como punto de salida y la norma del máximo, obtenemos $\delta = 0.7816\ldots$, $\theta = 1.3335\ldots$ y $C = 0.1853\ldots$ Ahora, vemos que la segunda condición de (4.4) del teorema 4.1 no se cumple porque $\ell\delta b^2 = 0.5295\ldots > \dfrac{1}{2}$. En consecuencia, no podemos garantizar la convergencia del método de Steffensen mediante la teoría de Kantorovich con el terorema 4.1 del capítulo anterior.

Sin embargo, sí que la podemos garantizar mediante el teorema 5.2 de este capítulo, puesto que la ecuación (5.5), que se reduce a

$$\frac{(0.2272\ldots)(t - 2.5669\ldots)(t - 1.6860\ldots)}{(0.2272\ldots)t - (0.7135\ldots)} = 0,$$

tiene dos raíces reales positivas y la más pequeña, $R = 1.6860\ldots$, cumple la condicióon (5.6), ya que

$$M + \beta\omega(R, R + \delta) = 0.6695\ldots < 1,$$

donde $M = 0.1976\ldots$ y $\omega(s,t) = K(s+t)$ con $K = 0.0926\ldots$

En la situación anterior, utilizando el método de Steffensen, después de cinco iteraciones y usando el criterio de parada $\|\mathbf{x_n} - \mathbf{x_{n-1}}\| < 10^{-16}$, obtenemos la aproximación numérica $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_8^*)^T$ de una solución de (5.17) que se puede ver en la tabla 5.1. En la tabla 5.2 mostramos los errores $\|\mathbf{x_n} - \mathbf{x}^*\|$ obtenidos usando el mismo criterio de parada. Notemos que el vector dado en la tabla 5.1 es una buena aproximación de la solución del sistema (5.17), puesto que $\|F(\mathbf{x}^*)\| \leq \text{constante} \times 10^{-16}$. Mostramos la sucesión $\{\|F(\mathbf{x_n})\|\}$ en la tabla 5.2.

Además, por el teorema 5.2, la existencia y unicidad de la solución está garantizada en la bola $\overline{B(\mathbf{x_0}, 1.6860\ldots)}$.

i	x_i^*	i	x_i^*	i	x_i^*	i	x_i^*
1	$2.042845\ldots$	3	$2.422455\ldots$	5	$2.576941\ldots$	7	$2.206427\ldots$
2	$2.206427\ldots$	4	$2.576941\ldots$	6	$2.422455\ldots$	8	$2.042845\ldots$

Tabla 5.1: Aproximación de la solución $\mathbf{x}^*$ de (5.17)

n	$\|\mathbf{x_n} - \mathbf{x}^*\|$	$\|F(\mathbf{x_n})\|$
0	$1.1769\ldots$	$7.8163\ldots \times 10^{-1}$
1	$2.2469\ldots \times 10^{-1}$	$1.3937\ldots \times 10^{-1}$
2	$8.8921\ldots \times 10^{-3}$	$5.4331\ldots \times 10^{-3}$
3	$1.4371\ldots \times 10^{-5}$	$8.7842\ldots \times 10^{-6}$
4	$3.7504\ldots \times 10^{-11}$	$2.29284\ldots \times 10^{-11}$

Tabla 5.2: Errores absolutos y $\|F(\mathbf{x_n})\|$ para (5.17)

Sistema de ecuaciones no lineales no diferenciable

Si $\mathbf{f} = \mathbf{2} = (2, 2, \ldots, 2)^T$ y $\delta = \mu = \frac{1}{2}$, entonces el sistema (5.16) se reduce a

$$F(\mathbf{x}) \equiv \mathbf{x} - \mathbf{2} - \frac{1}{2}A(\hat{\mathbf{x}} + \tilde{\mathbf{x}}) = 0, \qquad F : \mathbb{R}^8 \longrightarrow \mathbb{R}^8. \tag{5.18}$$

En este caso, $[\mathbf{u}, \mathbf{v}; F] = I - \frac{1}{2}A(\text{diag}\{\mathbf{z}\} + \text{diag}\{\mathbf{w}\})$ y el sistema de ecuaciones no lineales es no diferenciable ($\mu \neq 0$).

Si elegimos $\mathbf{x_0} = \left(\frac{12}{5}, \frac{12}{5}, \ldots, \frac{12}{5}\right)^T$ como punto de salida y la norma del máximo, obtenemos $\delta = 0.3594\ldots$, $\beta = 1.5115\ldots$ y la ecuación (5.5), que se reduce a

$$\frac{(0.1867\ldots)(t - 2.2255\ldots)(t - 1.1385\ldots)}{(0.1867\ldots)t - (0.5085\ldots)} = 0,$$

tiene dos raíces reales positivas y la más pequeña, $R = 1.1385\ldots$, cumple la condición (5.6), ya que

$$M + \beta\omega(R, R + \delta) = 0.7040\ldots < 1,$$

donde $M = 0.2710\ldots$ y $\omega(s, t) = L + K(s + t)$ con $L = 0.1235\ldots$ y $K = 0.0617\ldots$

A continuación, utilizando el método de Steffensen, después de cuatro iteraciones y usando el criterio de parada $\|\mathbf{x_n} - \mathbf{x_{n-1}}\| < 10^{-16}$, obtenemos la aproximación numérica $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_8^*)^T$ de una solución de (5.18) que se puede ver en la tabla 5.3. En la tabla 5.4 mostramos los errores $\|\mathbf{x_n} - \mathbf{x}^*\|$ obtenidos usando el mismo criterio de parada. Notemos que el vector dado en la tabla 5.3 es una buena aproximación de la solución del sistema (5.18), puesto que $\|F(\mathbf{x}^*)\| \leq \text{constante} \times 10^{-16}$. Mostramos la sucesión $\{\|F(\mathbf{x_n})\|\}$ en la tabla 5.4.

Además, por el teorema 5.2, la existencia y unicidad de la solución está garantizada en la bola $\overline{B(\mathbf{x_0}, 1.1385\ldots)}$.

i	x_i^*	i	x_i^*	i	x_i^*	i	x_i^*
1	2.039360...	3	2.383586...	5	2.521279...	7	2.188708...
2	2.188708...	4	2.521279...	6	2.383586...	8	2.039360...

Tabla 5.3: Aproximación de la solución $\mathbf{x}^*$ de (5.18)

n	$\|\mathbf{x_n} - \mathbf{x}^*\|$	$\|F(\mathbf{x_n})\|$
0	$3.6063\ldots \times 10^{-1}$	$3.5949\ldots \times 10^{-1}$
1	$2.3907\ldots \times 10^{-3}$	$1.6324\ldots \times 10^{-3}$
2	$6.7249\ldots \times 10^{-7}$	$4.6622\ldots \times 10^{-7}$
3	$5.0981\ldots \times 10^{-14}$	$3.5423\ldots \times 10^{-14}$

Tabla 5.4: Errores absolutos y $\|F(\mathbf{x_n})\|$ para (5.18)

5.2. Método predictor: el método simplificado de Steffensen

A la hora de mejorar la accesibilidad del método de Steffensen, proponemos, en este capítulo, un método iterativo híbrido (predictor-corrector) que no utiliza derivadas en su algoritmo. Teniendo en cuenta que el resultado de convergencia semilocal que acabamos de establecer para el método de Steffensen, teorema 5.2, es aplicable a operadores no diferenciables, consideramos, como método predictor para el método híbrido, un método iterativo que no utilice derivadas. En concreto, utilizaremos el método simplificado de Steffensen (5.4), que, como veremos, tiene mayor dominio de parámetros que el método de Steffensen.

5.2.1. Convergencia semilocal

Comenzamos estudiando la convergencia semilocal del método simplificado de Steffensen (5.4) bajo las mismas hipótesis generales que para el método de Steffensen, aunque, en este caso, como el objetivo de la sección es comparar los dominios de parámetros de los métodos simplificado de Steffensen (5.4) y de Steffensen, tanto en situaciones diferenciables como no diferenciables, exigimos directamente la condición (H_3) en vez de (C_3). Así, suponemos que se cumplen:

- (H_1) $\|F(z_0)\| \leq \delta_0$,
- (H_2) existe $[z_0, z_0 + F(z_0); F]^{-1} = B_0^{-1} \in \mathcal{L}(X, X)$, para $z_0 \in \Omega$, y es tal que $\|B_0^{-1}\| \leq \beta_0$,
- (H_3) $\|[x, y; F] - [u, v; F]\| \leq L + K(\|x - u\| + \|y - v\|); L, K \geq 0; x, y, u, v \in \Omega; x \neq y; u \neq v$.

Ahora, damos el siguiente lema técnico que utilizamos después.

Lema 5.4. *Sea $\{z_n\}$ la sucesión dada por el método simplificado de Steffensen (5.4). Si $z_{m-1} \neq z_m$ con $z_{m-1}, z_m \in \Omega$, entonces*

$$F(z_m) = ([z_m, z_{m-1}; F] - B_0)(z_m - z_{m-1}).$$

Demostración. A partir de la definición de $\{z_n\}$, se sigue

$$F(z_{m-1}) + B_0(z_m - z_{m-1}) = 0,$$

de manera que

$$\begin{aligned} F(z_m) &= F(z_m) - F(z_{m-1}) - B_0(z_m - z_{m-1}) \\ &= [z_m, z_{m-1}; F](z_m - z_{m-1}) - B_0(z_m - z_{m-1}) \\ &= ([z_m, z_{m-1}; F] - B_0)(z_m - z_{m-1}). \quad \blacksquare \end{aligned}$$

A continuación presentamos un resultado de convergencia semilocal para el método simplificado de Steffensen (5.4).

Teorema 5.5. *Sean X un espacio de Banach y $F : \Omega \subset X \to X$ un operador no lineal definido en un dominio abierto convexo no vacío Ω. Suponemos que se cumplen las condiciones (H_1)–(H_3). Si la ecuación*

$$t = \left(\frac{1 + N - \beta_0(L + K(2t + \delta_0))}{1 - \beta_0(L + K(2t + \delta_0))} \right) \beta_0 \delta_0, \tag{5.19}$$

donde $N = \beta_0(L + K\delta_0(1 + \beta_0))$, tiene al menos una raíz real positiva, y denotamos por r la raíz positiva más pequeña de (5.19),

$$Q = \beta_0(L + K(2r + \delta_0)) < 1 \tag{5.20}$$

y $B(z_0, r) \subset \Omega$, entonces el método simplificado de Steffensen (5.4), empezando en z_0, está bien definido y converge a una solución z^* de $F(x) = 0$. Además, la solución z^* y las iteraciones z_n pertenecen a $\overline{B(x_0, r)}$ y z^* es única en $\overline{B(x_0, r)} \cap \Omega$.

Demostración. Comenzamos probando que la sucesión $\{z_n\}$ está bien definida, es decir, $z_n \in B(z_0, r) \subset \Omega$, para todo $n \in \mathbb{N}$. Notemos que la raíz real positiva más pequeña r de la ecuación (5.19) es:

$$r = \left(1 + \frac{N}{1 - Q} \right) \beta_0 \delta_0. \tag{5.21}$$

A partir de (H_1)–(H_2) y como consecuencia de (5.21), se sigue que z_1 está bien definida y

$$\|z_1 - z_0\| \leq \|B_0^{-1}\| \|F(z_0)\| \leq \beta_0 \delta_0 < r.$$

Luego, $z_1 \in B(z_0, r)$.

A continuación, podemos definir $z_2 = z_1 - B_0^{-1} F(z_1)$ y

$$\begin{aligned}
\|z_2 - z_1\| &\leq \|B_0^{-1}\| \|F(z_1)\| \\
&\leq \|B_0^{-1}\| \|[z_1, z_0; F] - B_0\| \|z_1 - z_0\| \\
&\leq \beta_0 \left(L + K(\|z_1 - z_0\| + \|F(z_0)\|) \right) \|z_1 - z_0\| \\
&\leq \beta_0(L + K\delta_0(1 + \beta_0)) \|z_1 - z_0\| \\
&= N \|z_1 - z_0\|.
\end{aligned}$$

Además, por (5.21), también se tiene que

$$\|z_2 - z_0\| \leq \|z_2 - z_1\| + \|z_1 - z_0\| \leq (1 + N)\|z_1 - z_0\| \leq (1 + N)\beta_0 \delta_0 < r$$

y, por tanto, $z_2 \in B(z_0, r)$.

Suponemos ahora, para $i = 1, 2, \ldots, n$, que

$$\begin{aligned}
\|z_i - z_{i-1}\| &< N\, Q^{i-2} \|z_1 - z_0\|, \\
\|z_i - z_0\| &< \left(1 + N \, \frac{1 - Q^{i-1}}{1 - Q} \right) \|z_1 - z_0\| < r,
\end{aligned}$$

donde $Q < 1$ por (5.20).

Entonces, $z_{n+1} = z_n - B_0^{-1}F(z_n)$ está bien definido y

$$\begin{aligned}
\|z_{n+1} - z_n\| &\leq \|B_0^{-1}\|\|F(z_n)\| \\
&\leq \|B_0^{-1}\|\|[z_n, z_{n-1}; F] - B_0\|\|z_n - z_{n-1}\| \\
&\leq \beta_0 \left(L + K(\|z_n - z_0\| + \|z_{n-1} - z_0\| + \|F(z_0)\|)\right) \|z_n - z_{n-1}\| \\
&< \beta_0(L + K(2r + \delta_0))\|z_n - z_{n-1}\| \\
&= Q\|z_n - z_{n-1}\| \\
&< N\, Q^{n-1}\|z_1 - z_0\|.
\end{aligned}$$

Notemos que las diferencias divididas de primer orden $[z_n, z_{n-1}; F]$ existen ya que si $z_n = z_{n-1}$, z_{n-1} es una solución de $F(x) = 0$, la sucesión $\{z_n\}$ sería convergente y el resultado quedaría probado.

Además, por (5.20), también tenemos

$$\begin{aligned}
\|z_{n+1} - z_0\| &\leq \|z_{n+1} - z_n\| + \|z_n - z_0\| \\
&< \left(N(Q^{n-1} + \cdots + Q + 1) + 1\right) \|z_1 - z_0\| \\
&= \left(1 + N\,\frac{1 - Q^n}{1 - Q}\right) \|z_1 - z_0\| \\
&< \left(1 + \frac{N}{1 - Q}\right) \|z_1 - z_0\| \\
&\leq \left(1 + \frac{N}{1 - Q}\right) \beta_0\delta_0 \\
&= r.
\end{aligned}$$

Luego, $z_{n+1} \in B(z_0, r)$, para todo $n \in \mathbb{N}$.

Veamos ahora que la sucesión $\{z_n\}$ es de Cauchy. Como

$$\begin{aligned}
\|z_{n+j} - z_n\| &\leq \|z_{n+j} - z_{n+j-1}\| + \|z_{n+j-1} - z_{n+j-2}\| + \cdots + \|z_{n+1} - z_n\| \\
&< (Q^{j-1} + Q^{j-2} + \cdots + Q + 1)\|z_{n+1} - z_n\| \\
&= \frac{1 - Q^j}{1 - Q}\|z_{n+1} - z_n\| \\
&< \frac{1 - Q^j}{1 - Q} N\, Q^{n-1}\|z_1 - z_0\| \\
&< N\frac{Q^{n-1}}{1 - Q}\|z_1 - z_0\|,
\end{aligned}$$

para $j \geq 1$ y $Q < 1$, se sigue que $\{z_n\}$ es una sucesión de Cauchy. Por tanto, $\{z_n\}$ es convergente. Ahora, si $\lim_n z_n = z^* \in \overline{B(z_0, r)}$, vemos que z^* es una solución de $F(x) = 0$. Como

$$\|F(z_n)\| < (L + K(2r + \delta_0))\|z_n - z_{n-1}\|$$

y $\|z_n - z_{n-1}\| \to 0$, cuando $n \to \infty$, se sigue fácilmente, por la continuidad del operador F, que $F(z^*) = 0$.

Terminamos probando la unicidad de la solución z^* en $\overline{B(z_0, r)}$. Suponemos entonces que existe otra solución $y^* \in \overline{B(z_0, r)}$, con $y^* \neq z^*$, de la ecuación $F(x) = 0$. Sea $J = [y^*, z^*; F]$. Si J es inversible, tenemos $z^* = y^*$, puesto que $J(y^* - z^*) = F(y^*) - F(z^*)$. Para ver que J es inversible, basta, por el lema de Banach (lema 1.22), con ver que $\|I - B_0^{-1}J\| < 1$. En efecto, por hipótesis,

$$\begin{aligned}
\|I - B_0^{-1}J\| &\leq \|B_0^{-1}\|\|B_0 - J\| \leq \|B_0^{-1}\|\|[z_0, z_0 + F(z_0); F] - [y^*, z^*; F]\| \\
&\leq \beta_0(L + K(\|y^* - z_0\| + \|z^* - z_0 - F(z_0)\|)) \\
&\leq \beta_0(L + K(\|y^* - z_0\| + \|z^* - z_0\| + \|F(z_0)\|)) \\
&\leq \beta_0(L + K(2r + \delta_0)) \\
&< 1.
\end{aligned}$$

Luego, el operador J^{-1} existe. ■

5.2.2. Accesibilidad

Una vez probada la convergencia semilocal del método simplificado de Steffensen (5.4), nuestro siguiente objetivo es ver cuál es el dominio de parámetros de este método para compararlo con el de Steffensen. Para ello, transformamos la ecuación (5.19) en la siguiente ecuación cuadrática:

$$2K\beta_0 t^2 + (\beta_0(L + K\delta_0(1 - 2\beta_0)) - 1)\, t + \delta_0\beta_0(1 + K\delta_0\beta_0^2) = 0. \tag{5.22}$$

Como el término independiente de la ecuación anterior es siempre positivo, para que dicha ecuación tenga al menos una raíz real positiva, las dos raíces tienen que ser positivas. Vemos entonces cuándo tiene dos raíces reales positivas, lo que ocurre si

$$\beta_0(L + K\delta_0(1 - 2\beta_0)) - 1 < 0 \tag{5.23}$$

y

$$\begin{aligned}
\widetilde{\Delta} &= (\beta_0(L + K\delta_0(1 - 2\beta_0)) - 1)^2 - 8K\delta_0\beta_0^2(1 + K\delta_0\beta_0^2) \\
&= \left(\beta_0(L + K\delta_0(1 - 2\beta_0)) - 1 + \sqrt{8K\delta_0\beta_0^2(1 + K\delta_0\beta_0^2)}\right) \\
&\quad \times \left(\beta_0(L + K\delta_0(1 - 2\beta_0)) - 1 - \sqrt{8K\delta_0\beta_0^2(1 + K\delta_0\beta_0^2)}\right) > 0.
\end{aligned}$$

Observamos que los dos factores de $\widetilde{\Delta}$ son < 0 si

$$\beta_0(L + K\delta_0(1 - 2\beta_0)) + \sqrt{8K\delta_0\beta_0^2(1 + K\delta_0\beta_0^2)} < 1. \tag{5.24}$$

Luego, $\widetilde{\Delta} > 0$ si se cumple (5.24). También, como consecuencia de (5.24), se cumple (5.23). En este caso, la raíz positiva más pequeña de (5.22) es

$$r = \frac{1}{4K\beta_0}\left(1 - \beta_0(L + K\delta_0(1 - 2\beta_0)) - \sqrt{\widetilde{\Delta}}\right). \tag{5.25}$$

A continuación, vemos cuándo se cumple la condición (5.20) del teorema 5.5. Para ello, sustituimos el valor de r en (5.20) y obtenemos

$$1 - \beta_0(L + K\delta_0(1 + 2\beta_0) + \sqrt{\widetilde{\Delta}} > 0. \tag{5.26}$$

Notemos que también podemos considerar que la ecuación (5.22) tenga una raíz real positiva doble sin más que tener en cuenta la desigualdad no estricta en (5.24).

En consecuencia, las condiciones de convergencia que se imponen a los parámetros δ_0, β_0, L y K, como son que la ecuación (5.19) tenga al menos una raíz real positiva y que la raíz real positiva más pequeña de (5.19), denotada por r, cumpla (5.20), se van a cumplir siempre que se cumplan (5.24) y (5.26). Así, enunciamos entonces el siguiente resultado, cuya demostración se sigue fácilmente sin más que satisfacer las hipótesis del teorema 5.5.

Corolario 5.6. *Sean X un espacio de Banach y $F : \Omega \subset X \to X$ un operador no lineal definido en un dominio abierto convexo no vacío Ω. Supongamos que se cumplen las condiciones (H_1)–(H_3). Además, si se cumplen (5.24), (5.26) y $B(z_0, r) \subset \Omega$, donde r está definido en (5.25), entonces el método simplificado de Steffensen (5.4), empezando en z_0, está bien definido y converge a una solución z^* de $F(x) = 0$. Además, la solución z^* y las iteraciones z_n pertenecen a $\overline{B(z_0, r)}$ y z^* es única en $\overline{B(z_0, r)} \cap \Omega$.*

A continuación, representamos gráficamente el dominio de parámetros asociado al corolario 5.6. Para ello, seguimos el mismo criterio que para el corolario 5.3 y distinguimos dos casos: el caso diferenciable ($L = 0$) y el no diferenciable ($L \neq 0$). Notemos que las condiciones iniciales (H_1) y (H_2), exigidas al punto de salida z_0, definen los parámetros δ_0 y β_0, mientras que la condición (H_3), exigida al operador F, define los parámetros fijos L y K. Igual que antes, consideramos $x = \beta_0$ (eje de abscisas), $y = K\delta_0$ (eje de ordenadas) y coloreamos en el plano xy los valores de los parámetros que verifican las condiciones (5.24) y (5.26) del corolario 5.6. Observamos entonces que los ejes de los dominios de parámetros asociados a los corolarios 5.3 y 5.6 representan los mismos valores. A la vista de las figuras 5.3 y 5.4, podemos decir que el dominio de parámetros es mayor cuanto menor es el valor de L. Notamos que las regiones están superpuestas.

Si ahora comparamos los dominios de parámetros de los métodos de Steffensen y simplificado de Steffensen (5.4), figuras 5.5 y 5.6, vemos en ambos casos que el dominio de parámetros del método simplificado de Steffensen (5.4) es mayor que el del método de Steffensen. Notamos que las regiones están superpuestas.

5.3. Método iterativo híbrido (predictor-corrector)

A la vista de todas las figuras anteriores acerca de los dominios de parámetros de los métodos de Steffensen y simplificado de Steffensen (5.4), observamos que tenemos más posibilidades de localizar puntos de salida para obtener convergencia semilocal del método simplificado de Steffense (5.4)n que para el método de Steffensen. Por tanto, las condiciones que garantizan la convergencia semilocal del método de Steffensen son más restrictivas que las que garantizan la del método simplificado de Steffensen (5.4). Así, construimos un método iterativo híbrido (predictor-corrector), donde el método predictor es el método simplificado de Steffensen (5.4) y el método corrector es el método de Steffensen, que mejora la aplicabilidad del método de Steffensen.

5.3.1. Construcción del método

Nuestro objetivo inmediato consiste en asegurar que para un par (δ_0, β_0) que satisfaga las condiciones del corolario 5.6, es decir, que (δ_0, β_0) esté dentro del dominio de parámetros

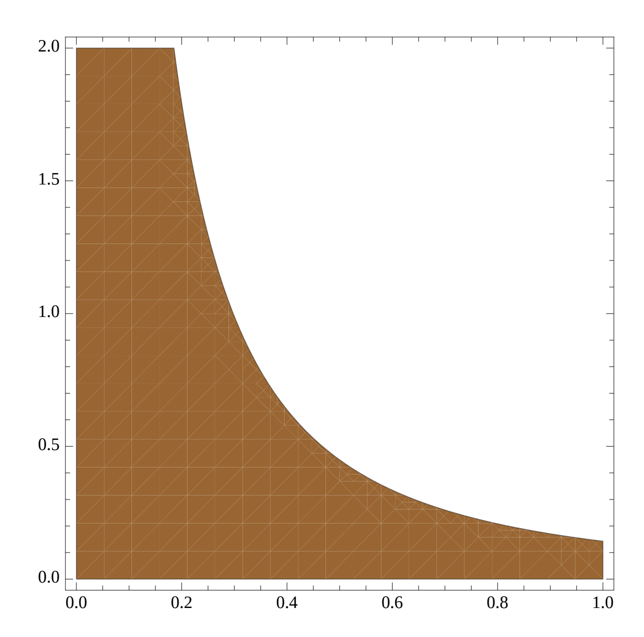


Figura 5.3: Dominio de parámetros del método simplificado de Steffensen asociado al corolario 5.6 cuando $L = 0$.

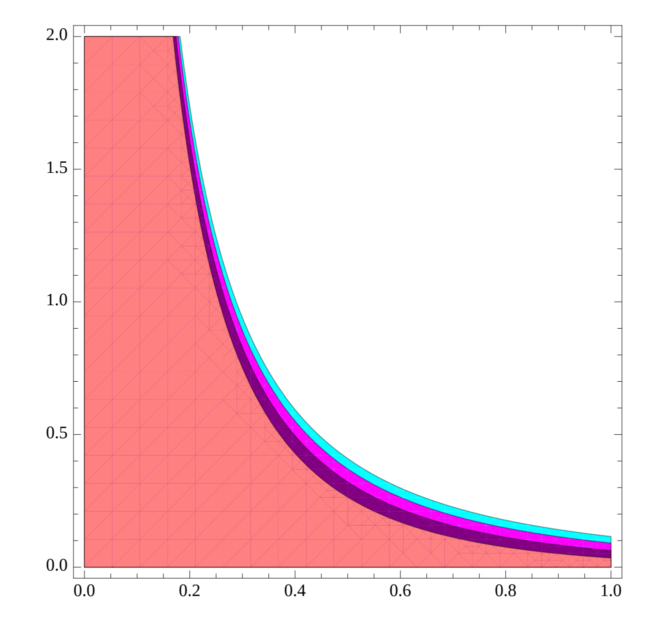


Figura 5.4: Dominios de parámetros del método simplificado de Steffensen asociados al corolario 5.6 cuando $L = \frac{1}{2}, \frac{1}{3}, \frac{1}{5}, \frac{1}{10}$ (rosa, morado, magenta y cyan, respectivamente).

del método simplificado de Steffensen (5.4), obtengamos un par $(\delta_{N_0}, \beta_{N_0})$ que satisfaga las condiciones del corolario 5.3, después de realizar un cierto número N_0 de iteraciones con el método simplificado de Steffensen (5.4), y asegurar así la convergencia del método de Steffensen al empezar este método en la iteración N_0 obtenida previamente mediante el método simplificado de Steffensen (5.4). Cuando esto ocurra, podemos considerar el par $(\delta_{N_0}, \beta_{N_0})$ como par inicial (δ, β) para el método de Steffensen.

Para ello, construimos una sencilla modificación del método de Steffensen que sea convergente cuando se tomen como puntos de salida los mismos que, a partir de los cuales, garantizan la convergencia del método simplificado de Steffensen (5.4). Así, consideramos el siguiente método iterativo híbrido (predictor-corrector):

$$\left\{ \begin{array}{l} \left\{ \begin{array}{l} \text{dado } z_0 \text{ en } \Omega, \\ z_{j+1} = z_j - [z_0, z_0 + F(z_0); F]^{-1} F(z_j), \quad j = 0, 1, \ldots, N_0 - 1, \end{array} \right. \\ \left\{ \begin{array}{l} x_0 = z_{N_0} \in \Omega, \\ x_{n+1} = x_n - [x_n, x_n + F(x_n); F]^{-1} F(x_n), \quad n \geq 0, \end{array} \right. \end{array} \right. \tag{5.27}$$

donde z_0 satisface las condiciones del corolario 5.6 y x_0 las del corolario 5.3.

Para que (5.27) sea convergente, nos planteamos entonces dos cuestiones:

1. Localizar un punto de salida z_0, a partir del cual el método predictor, el método simplificado de Steffensen (5.4), converja.

2. A partir de la convergencia del método predictor, garantizar la existencia de un valor $N_0 \in \mathbb{N}$ tal que z_{N_0} se pueda tomar como punto de salida para el método corrector, el método de Steffensen, y asegurar así después la convergencia de este método al empezar en $x_0 = z_{N_0}$.

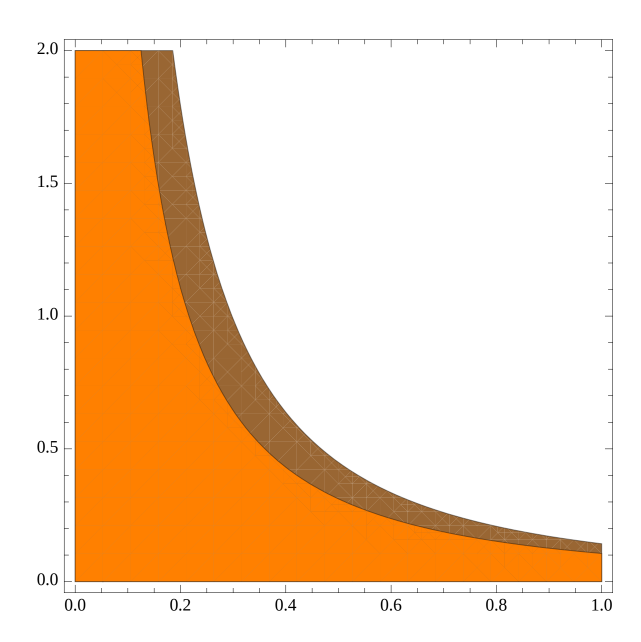


Figura 5.5: Dominios de parámetros de los métodos de Steffensen (anaranjado) y simplificado de Steffensen (marrón) cuando $L = 0$ (caso diferenciable).

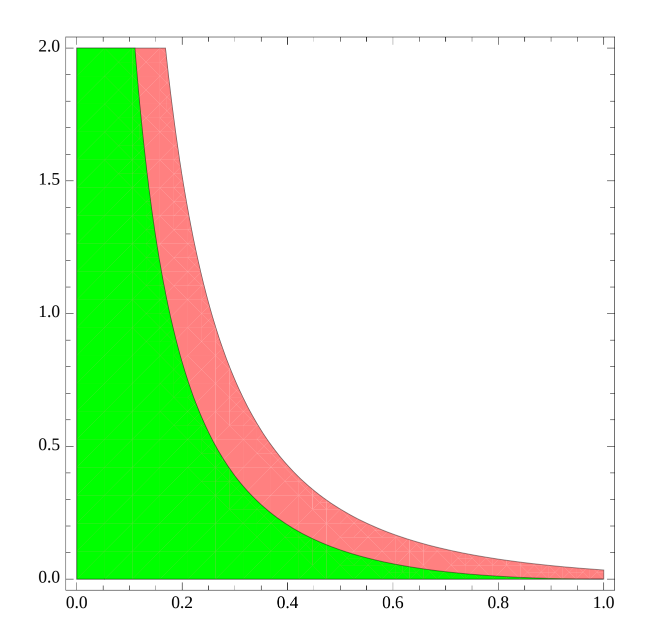


Figura 5.6: Dominios de parámetros de los métodos de Steffensen (verde) y simplificado de Steffensen (rosa) cuando $L = \frac{1}{2}$ (caso no diferenciable).

Entonces, utilizamos el método simplificado de Steffensen (5.4) durante un número finito de pasos N_0 hasta que $z_{N_0} = x_0$ cumpla las condiciones exigidas para que el método de Steffensen sea convergente y, después, aplicamos el método de Steffensen en vez del método simplificado de Steffensen (5.4). La clave del problema reside entonces en garantizar la existencia de N_0.

5.3.2. Convergencia semilocal del método

A continuación, estudiamos la convergencia semilocal del método híbrido (5.27). A partir del método predictor, el método simplificado de Steffensen (5.4), consideramos la siguiente situación. Dada la aproximación inicial z_0, consideramos la sucesión $\{z_n\}$ definida por el método simplificado de Steffensen (5.4) junto con

$$\|F(z_0)\| \leq \delta_0, \qquad \|[z_0, z_0 + F(z_0); F]^{-1}\| \leq \beta_0.$$

Para que el método simplificado de Steffensen (5.4) sea convergente, se tienen que cumplir las condiciones del corolario 5.6. Después, iterando, se van definiendo los pares (δ_n, β_n) asociados a cada z_n.

En primer lugar, observamos que la definición del par inicial (δ_0, β_0) para el método de Steffensen es inmediata porque los parámetros δ_0 y β_0 representan lo mismo en los métodos de Steffensen y simplificado de Steffensen (5.4). A continuación, procedemos de la siguiente forma.

Primer paso del método predictor: definición del par (δ_1, β_1).

Notemos

$$\begin{aligned}
\|F(z_1)\| &\leq \|[z_1, z_0; F] - B_0\|\|z_1 - z_0\| \\
&\leq (L + K(\|z_1 - z_0\| + \|F(z_0)\|))\,\|z_1 - z_0\| \\
&\leq (L + K(\beta_0\delta_0 + \delta_0))\,\beta_0\delta_0 \\
&< (L + K(2r + \delta_0))\,\beta_0\delta_0 \\
&= Q\delta_0 \\
&= \delta_1,
\end{aligned}$$

$$\begin{aligned}
\|I - B_0^{-1}A_1\| &\leq \|B_0^{-1}\|\|B_0 - A_1\| \\
&\leq \beta_0\,(L + K(\|z_1 - z_0\| + \|z_1 - z_0\| + \|F(z_1)\| + \|F(z_0)\|)) \\
&\leq \beta_0\,(L + K(2r + \delta_1 + \delta_0)) \\
&= \beta_0\,(L + K(2r + (1 + Q)\delta_0)) \\
&= T,
\end{aligned}$$

$$\|A_1^{-1}\| = \|[z_1, z_1 + F(z_1); F]^{-1}\| \leq \frac{\beta_0}{1 - T} = \beta_1,$$

siempre que $T < 1$, que puede escribirse como

$$T = (1 + K\delta_0\beta_0)Q < 1. \tag{5.28}$$

Además, $\delta_1 = Q\delta_0 < \delta_0$, puesto que $Q < 1$ si se cumple (5.28).

Segundo paso del método predictor: definición del par (δ_2, β_2).

Notemos

$$\begin{aligned}
\|F(z_2)\| &\leq \|[z_2, z_1; F] - B_0\|\|z_2 - z_1\| \\
&\leq (L + K(\|z_2 - z_0\| + \|z_1 - z_0 - F(z_0)\|))\,\|z_2 - z_1\| \\
&\leq (L + K(\|z_2 - z_0\| + \|z_1 - z_0\| + \|F(z_0)\|))\,\|B_0^{-1}\|\|F(z_1)\| \\
&\leq (L + K(2r + \delta_0))\,\beta_0\delta_1 \\
&= Q\delta_1 \\
&= \delta_2,
\end{aligned}$$

$$\begin{aligned}
\|I - B_0^{-1}A_2\| &\leq \|B_0^{-1}\|\|B_0 - A_2\| \\
&\leq \beta_0\,(L + K(\|z_2 - z_0\| + \|z_2 - z_0\| + \|F(z_2)\| + \|F(z_0)\|)) \\
&\leq \beta_0\,(L + K(2\|z_2 - z_0\| + \delta_2 + \delta_0)) \\
&= \beta_0\,(L + K(2r + (1 + Q^2)\delta_0)) \\
&< \beta_0\,(L + K(2r + (1 + Q)\delta_0)) \\
&= T,
\end{aligned}$$

$$\|A_2^{-1}\| = \|[z_2, z_2 + F(z_2); F]^{-1}\| \leq \frac{\beta_0}{1 - T} = \beta_2,$$

siempre que $T < 1$. Además, $\delta_2 = Q^2\delta_0 < \delta_0$, puesto que $Q < 1$ si se cumple (5.28).

***n*-ésimo paso del método predictor:** definición del par (δ_n, β_n).

Notemos

$$\begin{aligned}
\|F(z_n)\| &\leq \|[z_n, z_{n-1}; F] - B_0\|\|z_n - z_{n-1}\| \\
&\leq (L + K(\|z_{n-1} - z_0\| + \|z_{n-1} - z_0 - F(z_0)\|))\,\|z_n - z_{n-1}\| \\
&\leq (L + K(\|z_n - z_0\| + \|z_{n-1} - z_0\| + \|F(z_0)\|))\,\|B_0^{-1}\|\|F(z_{n-1})\| \\
&\leq (L + K(2r + \delta_0))\,\beta_0\delta_{n-1} \\
&= Q\delta_{n-1} \\
&= \delta_n,
\end{aligned}$$

$$\begin{aligned}
\|I - B_0^{-1}A_n\| &\leq \|B_0^{-1}\|\|B_0 - A_n\| \\
&\leq \beta_0\,(L + K(\|z_n - z_0\| + \|z_n - z_0\| + \|F(z_n)\| + \|F(z_0)\|)) \\
&\leq \beta_0\,(L + K(2r + \delta_n + \delta_0)) \\
&= \beta_0\,(L + K(2r + (1 + Q^n)\delta_0)) \\
&< \beta_0\,(L + K(2r + (1 + Q)\delta_0)) \\
&= T,
\end{aligned}$$

$$\|A_n^{-1}\| = \|[z_n, z_n + F(z_n); F]^{-1}\| \leq \frac{\beta_0}{1 - T} = \beta_n,$$

siempre que $T < 1$. Además, $\delta_n = Q^n\delta_0 < \delta_0$, puesto que $Q < 1$ si se cumple (5.28).

Una vez construido el par (δ_n, β_n), a partir de las sucesiones $\{\delta_n\}$ y $\{\beta_n\}$, tenemos que asegurar la existencia de un $N_0 \in \mathbb{N}$ tal que el par $(\delta_{N_0}, \beta_{N_0})$ satisfaga las condiciones de convergencia exigidas al método de Steffensen en el corolario 5.3. Si observamos las condiciones (5.12), (5.13) y (5.15) del corolario 5.3 y tenemos en cuenta $\delta_n < \delta_{n-1}$ y que $\beta_n = \frac{\beta_0}{1-T}$ es constante, para todo $n \in \mathbb{N}$, por ser $Q < 1$ y $T < 1$, podemos asegurar entonces las siguientes tres afirmaciones.

Primero, siempre que

$$L\beta < \frac{1 + L}{1 + 2L}, \tag{5.29}$$

donde $\beta = \frac{\beta_0}{1-T}$, existirá $N_1 \in \mathbb{N}$ para el que se cumpla la condición (5.12) del corolario 5.3; es decir,

$$\delta_{N_1}\left(M_{N_1}^2 + (L\beta_{N_1} + K\delta_{N_1}\beta_{N_1} - 1)\,(M_{N_1} + \beta_{N_1})\right) < 0,$$

donde $M_{N_1} = \beta_{N_1}(L + K\delta_{N_1}(1 + \beta_{N_1}))$.

Segundo, siempre que

$$L\beta < \frac{1}{2}, \tag{5.30}$$

donde $\beta = \frac{\beta_0}{1-T}$, existirá $N_2 \in \mathbb{N}$ para el que se cumpla la condición (5.13) del corolario 5.3; es decir,

$$\begin{aligned}
&M_{N_2}(1 - 2K\delta_{N_2}\beta_{N_2}) + \beta_{N_2}(L + K\delta_{N_2}(1 - 2\beta_{N_2})) \\
&\quad + \sqrt{-8K\delta_{N_2}\beta_{N_2}\left(M_{N_2}^2 + (L\beta_{N_2} + K\delta_{N_2}\beta_{N_2} - 1)\,(M_{N_2} + \beta_{N_2})\right)} < 1.
\end{aligned}$$

donde $M_{N_2} = \beta_{N_2}(L + K\delta_{N_2}(1+\beta_{N_2}))$.

Tercero, siempre que se cumpla (5.30), existirá $N_3 \in \mathbb{N}$ para el que se cumpla la condición (5.15) del corolario 5.3; es decir,

$$1 - L\beta_{N_3} - (1 + 2K\delta_{N_3}\beta_{N_3})M_{N_3} - K\delta_{N_3}\beta_{N_3}(1+2\beta_{N_3}) + \sqrt{\Delta_{N_3}} > 0.$$

donde $M_{N_3} = \beta_{N_3}(L + K\delta_{N_3}(1+\beta_{N_3}))$ y

$$\begin{aligned} \Delta_{N_3} &= (M_{N_3}(1 - 2K\delta_{N_3}\beta_{N_3}) + \beta_{N_3}(L + K\delta_{N_3}(1-2\beta_{N_3})) - 1)^2 \\ &\quad +8K\delta_{N_3}\beta_{N_3}\left(M_{N_3}^2 + (L\beta_{N_3} + K\delta_{N_3}\beta_{N_3} - 1)(M_{N_3} + \beta_{N_3})\right). \end{aligned}$$

Obsérvese que (5.29) se cumple siempre que se cumple (5.30). En consecuencia, siempre que se cumpla (5.30), podemos tomar $N_0 = \text{máx}\{N_1, N_2, N_3\}$, elegir $x_0 = z_{N_0}$ y aplicar el método de Steffensen a partir de la aproximación $x_0 = z_{N_0}$, dada por el método simplificado de Steffensen (5.4), para garantizar la convergencia del método híbrido (5.27).

Para terminar, resumimos todo lo anterior en el siguiente resultado.

Teorema 5.7. *Sean X un espacio de Banach y $F : \Omega \subset X \to X$ un operador no lineal definido en un dominio abierto convexo no vacío Ω. Supongamos que se cumplen las condiciones (H_1)–(H_3), (5.28), (5.30) y $B(z_0, r) \subset \Omega$, donde r está dado en (5.25). Entonces, existe un $N_0 \in \mathbb{N}$ tal que, para $x_0 = z_{N_0}$, la sucesión $\{x_n\}$ dada por el método híbrido (5.27) está bien definida y converge a una solución de $F(x) = 0$.*

5.4. Aplicación

Ilustramos el estudio realizado anteriormente con dos aplicaciones a sistemas no lineales, siendo uno diferenciable y otro no diferenciable. Ambos sistemas surgen de las discretizaciones de los problemas conservativos definidos en la sección 1.6.2 del capítulo 1.

5.4.1. Sistema de ecuaciones no lineales diferenciable

La distribución constante de la temperatura se conoce en una varilla homogénea de longitud 1 en la que, como consecuencia de una reacción química o algún proceso de calor, el calor se genera a una velocidad $\phi(x(t))$ por unidad de tiempo y por unidad de longitud, siendo $\phi(x(t))$ una función dada por el exceso de temperatura x de la varilla sobre la temperatura circulante. Consideramos el problema en el que al final de la varilla, $t = 0$ y $t = 1$, se mantienen esas temperaturas, discretizamos el correspondiente problema de contorno dado por (1.49)–(1.50) y aproximamos una solución del sistema no lineal (1.52) que surge del proceso de discretización.

Por ejemplo, si elegimos la ley exponencial, $\phi(u) = \exp(u)$, para la generación del calor y $m = 8$, entonces el vector $v_{\mathbf{x}}$ del correspondiente sistema (1.52) está dado por

$$v_{\mathbf{x}} = (v_1, v_2, \ldots, v_8)^t, \qquad v_i = \exp(x_i), \qquad i = 1, 2, \ldots, 8. \tag{5.31}$$

A continuación, observamos que una solución $\mathbf{x}^*$ del correspondiente sistema (1.52) con (5.31) satisface

$$\|\mathbf{x}^*\| \leq h^2\|A^{-1}\|\,\|v_{\mathbf{x}^*}\| \qquad \Rightarrow \qquad \|\mathbf{x}^*\| - h^2\|A^{-1}\|\exp(\|\mathbf{x}^*\|) \leq 0,$$

donde $\|A^{-1}\| = 10$ y $h = \frac{1}{9}$, de forma que $\|\mathbf{x}^*\| \in [0, \varrho_1] \cup [\varrho_2, +\infty]$, siendo $\varrho_1 = 0.142342\ldots$ y $\varrho_2 = 3.279579\ldots$ las dos raíces reales positivas de la ecuación escalar $81t - 10\exp(t) = 0$. Por tanto, podemos considerar $F : \Omega \subset \mathbb{R}^8 \to \mathbb{R}^8$ con

$$\Omega = \{\mathbf{x} \in \mathbb{R}^8;\ \|\mathbf{x}\| < 3\},$$

puesto que $\varrho_1 < 3 < \varrho_2$.

Además, la primera derivada de la función F definida en (1.52) está dada por

$$F'(\mathbf{x}) = A + h^2 \mathrm{diag}\{v_{\mathbf{x}}\}.$$

En consecuencia,

$$F'(\mathbf{x}) - F'(\mathbf{y}) = h^2 \mathrm{diag}\{\mathbf{z}\},$$

donde $\mathbf{y} = (y_1, y_2, \ldots, y_8)^t$ y $\mathbf{z} = (\exp(x_1) - \exp(y_1), \exp(x_2) - \exp(y_2), \ldots, \exp(x_8) - \exp(y_8))^t$, y

$$\|F'(\mathbf{x}) - F'(\mathbf{y})\| \leq h^2 \max_{1 \leq i \leq 8} |\exp(c_i)|\, \|\mathbf{x} - \mathbf{y}\|,$$

con $\mathbf{c} = (c_1, c_2, \ldots, c_8)^t \in \Omega$ y $h = \frac{1}{9}$, de modo que

$$\|F'(\mathbf{x}) - F'(\mathbf{y})\| \leq \mathrm{e}^3 h^2 \|\mathbf{x} - \mathbf{y}\|. \tag{5.32}$$

Considerando la diferencia dividida ([37])

$$[\mathbf{x}, \mathbf{y}; F] = \int_0^1 F'\left(\tau\mathbf{x} + (1-\tau)\mathbf{y}\right) d\tau$$

y teniendo en cuenta

$$\int_0^1 \|\tau(\mathbf{x} - \mathbf{u}) + (1-\tau)(\mathbf{y} - \mathbf{v})\| d\tau \leq \frac{1}{2}\left(\|\mathbf{x} - \mathbf{u}\| + \|\mathbf{y} - \mathbf{v}\|\right),$$

y (5.32), tenemos

$$\begin{aligned} \|[\mathbf{x}, \mathbf{y}; F] - [\mathbf{u}, \mathbf{v}; F]\| &\leq \int_0^1 \|F'\left(\tau\mathbf{x} + (1-\tau)\mathbf{y}\right) - F'\left(\tau\mathbf{u} + (1-\tau)\mathbf{v}\right)\| \, d\tau \\ &\leq \mathrm{e}^3 h^2 \int_0^1 \left(\tau\|\mathbf{x} - \mathbf{u}\| + (1-\tau)\|\mathbf{y} - \mathbf{v}\|\right) d\tau \\ &= \frac{\mathrm{e}^3}{2} h^2 \left(\|\mathbf{x} - \mathbf{u}\| + \|\mathbf{y} - \mathbf{v}\|\right). \end{aligned}$$

Luego, $\omega(s,t) = \frac{\mathrm{e}^3}{2} h^2 (s+t) = \frac{\mathrm{e}^3}{162}(s+t)$.

Eligiendo el punto de salida $\mathbf{x_0} = (0, 0, \ldots, 0)^t$ y la norma del máximo, se obtiene $\delta = \frac{1}{81}$ y $\beta = 11.177516\ldots$, de manera que la ecuación (5.5) del teorema 5.2, que se reduce a

$$\frac{(2.771684\ldots)t^2 - (1.164148\ldots)t + (0.137625\ldots)}{(2.771684\ldots)t - (0.774543\ldots)} = 0,$$

no tiene raíces reales. Así que no podemos garantizar la convergencia del método de Steffensen a una solución de (1.52) con $\phi(u) = \exp(u)$.

En cambio, si que podemos aplicar el método simplificado de Steffensen (5.4) empezando en $\mathbf{z_0} = (0, 0, \ldots, 0)^t$. Como $\delta_0 = \frac{1}{81}$, $\beta_0 = 11.177516\ldots$, la ecuación (5.19) del teorema 5.5, que se reduce a

$$\frac{(2.771684\ldots)(t - 0.283165\ldots)(t - 0.209447\ldots)}{(2.771684\ldots)t - (0.982890\ldots)} = 0,$$

tiene dos raíces reales positivas y la más pequeña, $r = 0.209447\ldots$, cumple la condición (5.20), puesto que

$$Q = \beta_0(L + K(2r + \delta_0)) = 0.597631\ldots < 1,$$

y es tal que $B(\mathbf{z_0}, r) \subset \Omega = B(0, 3)$. Por tanto, el método simplificado de Steffensen (5.4) converge a una solución de (1.52) con $\phi(u) = \exp(u)$, la dada por el vector $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_8^*)^t$ que aparece en la tabla 5.5, después de ocho iteraciones con una tolerancia de 10^{-16}. Observamos que $\|\mathbf{x}^*\| = 0.138937\ldots < 3$ y $\mathbf{x}^*$ es única en la bola $\overline{B(\mathbf{z_0}, 0.209447\ldots)}$. En la tabla 5.6 se muestran los errores $\|\mathbf{z_n} - \mathbf{x}^*\|$ utilizando el criterio de parada $\|\mathbf{z_n} - \mathbf{z_{n-1}}\| < 10^{-16}$. Notemos que el vector dado en la tabla 5.5 es una buena aproximación de la solución del sistema (1.52) con $\phi(u) = \exp(u)$, puesto que $\|F(\mathbf{x}^*)\| \leq \text{constante} \times 10^{-16}$. Mostramos la sucesión $\{\|F(\mathbf{z_n})\|\}$ en la tabla 5.6.

i	x_i^*	n	x_i^*	i	x_i^*	n	x_i^*
1	$0.05481058\ldots$	3	$0.12475178\ldots$	5	$0.13893761\ldots$	7	$0.09657993\ldots$
2	$0.09657993\ldots$	4	$0.13893761\ldots$	6	$0.12475178\ldots$	8	$0.05481058\ldots$

Tabla 5.5: Aproximación de la solución $\mathbf{x}^*$ de (1.52) con $\phi(u) = \exp(u)$

n	$\|\mathbf{z_n} - \mathbf{x}^*\|$	$\|F(\mathbf{z_n})\|$
0	$1.3893\ldots \times 10^{-1}$	$1.2345\ldots \times 10^{-2}$
1	$9.4358\ldots \times 10^{-4}$	$1.1258\ldots \times 10^{-4}$
2	$1.3380\ldots \times 10^{-5}$	$1.6342\ldots \times 10^{-6}$
3	$1.9003\ldots \times 10^{-7}$	$2.3261\ldots \times 10^{-8}$
4	$2.6986\ldots \times 10^{-9}$	$3.3039\ldots \times 10^{-10}$
5	$3.8321\ldots \times 10^{-11}$	$4.6918\ldots \times 10^{-12}$
6	$5.4417\ldots \times 10^{-13}$	$6.6625\ldots \times 10^{-14}$
7	$7.7259\ldots \times 10^{-15}$	$9.4609\ldots \times 10^{-16}$

Tabla 5.6: Errores absolutos obtenidos con el método simplificado de Steffensen y $\{\|F(\mathbf{z_n})\|\}$

Por otra parte, es fácil ver que se puede aplicar el método híbrido (5.27) porque se cumplen las condiciones (5.28) y (5.30) del teorema 5.7, ya que

$$T = \beta_0 \left(L + K(2r + (1 + Q)\delta_0)\right) = 0.926161\ldots < 1 \qquad \text{y} \qquad L\beta = 0 < \frac{1}{2}$$

respectivamente, de manera que está garantizada la convergencia semilocal del método híbrido (5.27) para un cierto $N_0 \in \mathbb{N}$. Así, la primera aproximación dada por el método simplificado

de Steffensen ($N_0 = 1$) ya cumple las hipótesis del teorema 5.2, puesto que la raíz real positiva más pequeña de la ecuación (5.5), que ahora se reduce a

$$\frac{(2.812541\ldots)(t-0.354796\ldots)(t-0.001279\ldots)}{(2.812541\ldots)t-(0.997887\ldots)} = 0,$$

es $R = 0.001279\ldots$ y cumple (5.6):

$$M + \beta\omega(R, R+\delta) = 0.005711\ldots < 1.$$

Por tanto, después de una aproximación del método simplificado de Steffensen (5.4), podemos aplicar el método de Steffensen para aproximar la solución $\mathbf{x}^*$ dada en la tabla 5.5, que se obtiene después de dos aproximaciones más con el método de Steffensen y una tolerancia de 10^{-16}. En la tabla 5.7 se muestran los errores $\|\mathbf{x_n} - \mathbf{x}^*\|$ y la sucesión $\{\|F(\mathbf{x_n})\|\}$ utilizando el mismo criterio de parada que antes.

n	$\|\mathbf{x_n} - \mathbf{x}^*\|$	$\|F(\mathbf{x_n})\|$
0	$1.3893\ldots \times 10^{-1}$	$1.2345\ldots \times 10^{-2}$
1	$9.4358\ldots \times 10^{-4}$	$1.1258\ldots \times 10^{-4}$
2	$4.4714\ldots \times 10^{-8}$	$5.5579\ldots \times 10^{-9}$

Tabla 5.7: Errores absolutos obtenidos con el método híbrido (5.27) y $\{\|F(\mathbf{x_n})\|\}$

5.4.2. Sistema de ecuaciones no lineales no diferenciable

Recordemos primero que si consideramos un sistema de ecuaciones del tipo (1.52), donde la función $\phi(u)$ es no derivable, entonces la función F es no diferenciable.

Consideramos entonces $\phi(u) = \frac{u^2}{3} + |u| + 1$, de manera que el correspondiente sistema de ecuaciones no lineales dado por (1.52) con $m = 8$ y

$$v_{\mathbf{x}} = (v_1, v_2, \ldots, v_8)^t, \qquad v_i = \frac{x_i^2}{3} + |x_i| + 1, \qquad i = 1, 2, \ldots, 8, \tag{5.33}$$

es no diferenciable.

Como en $\mathbb{R}^8$ podemos considerar diferencias divididas de primer orden que no necesitan que la función sea diferenciable, consideramos la diferencia dividida de primer orden dada por $[\mathbf{u}, \mathbf{v}; F] = ([\mathbf{u}, \mathbf{v}; F]_{ij})_{i,j=1}^8 \in \mathcal{L}(\mathbb{R}^8, \mathbb{R}^8)$, donde $[\mathbf{u}, \mathbf{v}; F]_{ij}$ está definido en (1.46) si $m = 8$. Luego,

$$[\mathbf{u}, \mathbf{v}; F] = A + h^2 \operatorname{diag}\{\mathbf{z}\},$$

donde $z_i = \frac{u_i+v_i}{3} + \frac{|u_i|-|v_i|}{u_i-v_i}$, para $i = 1, 2, \ldots, 8$, y $h = \frac{1}{9}$. Así,

$$\begin{aligned}
\|[\mathbf{x}, \mathbf{y}; F] - [\mathbf{u}, \mathbf{v}; F]\| &\leq \max_{1\leq i\leq 8}\left|\frac{x_i-u_i}{3} + \frac{y_i-v_i}{3} + \frac{|x_i|-|y_i|}{x_i-y_i} - \frac{|u_i|-|v_i|}{u_i-v_i}\right| \\
&\leq h^2\left(\frac{1}{3}\max_{1\leq i\leq 8}|x_i-u_i+y_i-v_i| + \max_{1\leq i\leq 8}\left|\frac{|x_i|-|y_i|}{x_i-y_i} - \frac{|u_i|-|v_i|}{u_i-v_i}\right|\right) \\
&\leq \frac{h^2}{3}\left(\|\mathbf{x}-\mathbf{u}\| + \|\mathbf{y}-\mathbf{v}\|\right) + 2h^2,
\end{aligned}$$

de manera que $\omega(s,t) = 2h^2 + \frac{h^2}{3}(s+t) = \frac{2}{81} + \frac{1}{243}(s+t)$.

Si elegimos el punto de salida $\mathbf{x_0} = \left(\frac{1}{10}, \frac{1}{10}, \ldots, \frac{1}{10}\right)^T$ y la norma del máximo, obtenemos $\delta = 0.086378\ldots$ y $\beta = 11.260889\ldots$, de manera que la ecuación (5.5) del teorema 5.2, que se reduce a

$$\frac{(0.092682\ldots)t^2 - (0.483595\ldots)t + 0.709393\ldots}{(0.092682\ldots)t - (0.390824\ldots)} = 0,$$

no tiene raíces reales. Por lo tanto, no podemos garantizar la convergencia del método de Steffensen a una solución de (1.52) con (5.33).

En cambio, si que podemos aplicar el método simplificado de Steffensen (5.4) empezando en $\mathbf{z_0} = \left(\frac{1}{10}, \frac{1}{10}, \ldots, \frac{1}{10}\right)^T$. Como $\delta_0 = \delta$ y $\beta_0 = \beta$, la ecuación (5.19) del teorema 5.5, que se reduce a

$$\frac{(0.092682\ldots)(t - 7.194574\ldots)(t - 1.524492\ldots)}{(0.092682\ldots)t - (0.717950\ldots)} = 0,$$

tiene dos raíces reales positivas y la más pequeña, $r = 1.524492\ldots$, cumple la condición (5.20), puesto que

$$Q = \beta_0(L + K(2r + \delta_0)) = 0.423342\ldots < 1.$$

Por tanto, el método simplificado de Steffensen (5.4) converge a una solución de (1.52) con (5.33), la dada por el vector $\mathbf{x}^* = (x_1^*, x_2^*, \ldots, x_8^*)^t$ en la tabla 5.8, después de cinco iteraciones con una tolerancia de 10^{-16}. Observamos que $\mathbf{x}^*$ es única en la bola $\overline{B(\mathbf{z_0}, 1.524492\ldots)}$. En la tabla 5.9 mostramos los errores $\|\mathbf{z_n} - \mathbf{x}^*\|$ utilizando el criterio de parada $\|\mathbf{z_n} - \mathbf{z_{n-1}}\| < 10^{-16}$. Notemos que el vector dado en la tabla 5.8 es una buena aproximación de la solución del sistema (1.52) con $v_\mathbf{x}$ definido en (5.33), ya que $\|F(\mathbf{x}^*)\| \leq \text{constante} \times 10^{-16}$. Mostramos la sucesión $\{\|F(\mathbf{z_n})\|\}$ en la tabla 5.9.

i	x_i^*	n	x_i^*	i	x_i^*	n	x_i^*
1	0.05468713...	3	0.12442184...	5	0.13855711...	7	0.09634112...
2	0.09634112...	4	0.13855711...	6	0.12442184...	8	0.05468713...

Tabla 5.8: Aproximación de la solución $\mathbf{x}^*$ de (1.52) y (5.33)

n	$\|\mathbf{z_n} - \mathbf{x}^*\|$	$\|F(\mathbf{z_n})\|$
0	$4.5312\ldots \times 10^{-2}$	$8.6378\ldots \times 10^{-2}$
1	$1.3059\ldots \times 10^{-5}$	$7.6576\ldots \times 10^{-6}$
2	$1.8877\ldots \times 10^{-8}$	$3.4066\ldots \times 10^{-9}$
3	$2.7309\ldots \times 10^{-11}$	$4.9253\ldots \times 10^{-11}$
4	$3.9579\ldots \times 10^{-14}$	$7.1254\ldots \times 10^{-15}$

Tabla 5.9: Errores absolutos obtenidos con el método simplificado de Steffensen y $\{\|F(\mathbf{z_n})\|\}$

Por otra parte, es fácil ver que se puede aplicar el método híbrido (5.27) porque se cumplen las condiciones (5.28) y (5.30) del teorema 5.7, ya que

$$T = \beta_0 \left(L + K(2r + (1+Q)\delta_0)\right) = 0.425037\ldots < 1 \qquad \text{y} \qquad L\beta = 0.483590\ldots < \frac{1}{2}$$

respectivamente, de manera que está garantizada la convergencia semilocal del método híbrido (5.27) para un cierto $N_0 \in \mathbb{N}$. Así, la primera aproximación dada por el método simplificado de Steffensen ($N_0 = 1$) ya cumple las hipótesis del teorema 5.2, puesto que la raíz real positiva más pequeña de la ecuación (5.5), que ahora se reduce a

$$\frac{(0.069768\ldots)(t - 5.570018\ldots)(t - 3.459413\ldots)}{(0.069768\ldots)t - (0.508965\ldots)} = 0,$$

es $R = 0.000142\ldots$ y cumple (5.6):

$$M + \beta\omega(R, R + \delta) = 0.556922\ldots < 1.$$

Por tanto, después de una aproximación del método simplificado de Steffensen (5.4), podemos aplicar el método de Steffensen para aproximar la solución $\mathbf{x}^*$ dada en la tabla 5.5, que se obtiene después de dos aproximaciones más con el método de Steffensen y una tolerancia de 10^{-16}. En la tabla 5.10 se muestran los errores $\|\mathbf{x_n} - \mathbf{x}^*\|$ y la sucesión $\{\|F(\mathbf{x_n})\|\}$ utilizando el mismo criterio de parada que antes.

n	$\|\mathbf{x_n} - \mathbf{x}^*\|$	$\|F(\mathbf{x_n})\|$
0	$4.5312\ldots \times 10^{-2}$	$8.6378\ldots \times 10^{-2}$
1	$1.3059\ldots \times 10^{-5}$	$7.6576\ldots \times 10^{-6}$
2	$3.2146\ldots \times 10^{-12}$	$4.8940\ldots \times 10^{-13}$

Tabla 5.10: Errores absolutos obtenidos con el método híbrido (5.27) y $\{\|F(\mathbf{x_n})\|\}$

Bibliografía

[1] Amat, S., y Busquier, S., *Convergence and numerical analysis of a family of two-step Steffensen's methods*, Comput. Math. Appl., **49,2**: 13–22, 2005.

[2] Amat, S., y Busquier, S., *A two-step Steffensen's method under modified convergence conditions*, J. Math. Anal. Appl., **324,2**: 1084–1092, 2006.

[3] Andronow, A. A., y Chaikin, C. E., *Theory of oscillations*, Princenton University Press, New Jersey, 1949.

[4] Argyros, I. K., *On the Secant Method*, Publ. Math. Debrecen, **43**: 223–238, 1993.

[5] Argyros, I. K., *A new convergence theorem for Steffensen's method on Banach spaces and applications*, Southwest J. Pure Appl. Math., **1**: 23–29, 1997.

[6] Atkinson, K. E., *The numerical solution of a nonlinear boundary integral equation on smooth surfaces*, IMA J. of Numer. Anal., **14**: 461–483, 1994.

[7] Balazs, M., y Goldner, G., *On existence of divided differences in linear spaces*, Rev. Anal. Number. Theorie Approximation, **2**: 5–9, 1973.

[8] Banach, S., *Théorie des opérations linéaires*, Chelsea, 2ª edición, New York, 1963.

[9] Banás, J., Rocha Martin, C. J., y Sadarangani, K., *On solutions of a quadratic integral equation of Hammerstein type*, Math. Comput. Modelling, **43**: 97–104, 2006.

[10] Beauzamy, B., *Introduction to Banach spaces and their geometry*, North Holland, 1985.

[11] Berberian, S. K., *Lectures in functional analysis and operator theory*, Springer Verlag, 1974.

[12] Bruns, D. D., y Bailey, J. E., *Nonlinear feedback control for operating a nonisothermal CSTR near an unstable steady state*, Chem. Eng. Sci., **32**: 257–264, 1977.

[13] Cartan, H., *Calcul diffeérentiel*, Hermann, 1971.

[14] Corduneanu, C., *Integral Equations and Applications*, Cambridge University Press, Cambridge, 1991.

[15] Curtain, R. F., y Pritchard, A. J., *Functional analysis in modern applied mathematics*, Academic Press, 1977.

[16] Day, M. M., *Normed linear spaces*, Springer Verlag, 1958.

[17] Deimling, K., *Nonlinear functional analysis.* Springer Verlag, Berlin, 1985.

[18] Dennis, J. E., Jr., *Toward a Unified Convergence Theory for Newton-Like Methods.* Article in Nonlinear funcional analysis and aplications. Edited by L. B. Rall, Academic Press, New York, **43**: 425–472, 1970.

[19] Dunford, N., y Schwartz, J. T., *Linear operators. Part I: General theory*, Interscience Publishers Inc., 1958.

[20] Ezquerro, J. A., Gutiérrez, J. M., Hernández, M. A., Romero, N., y Rubio, M. J., *El método de Newton: de Newton a Kantorovich*, La Gaceta de la RSME, **13,1**: 53–76, 2010.

[21] Faraci, F., y Moroz, V., *Solutions of Hammerstein integral equations via a variational principle*, J. Integral Equations. Anal., **15,4**: 385–402, 2003.

[22] Ganesh, M., y Joshi, M. C., *Numerical solvability of Hammerstein integral equations of mixed type*, IMA J. Numer. Anal., **11**: 21–31, 1991.

[23] Hernández, M. A., y Rubio, M. J., *A uniparametric family of iterative processes for solving nondifferentiable equations*, J. Math. Anal. Appl., **275**: 821–834, 2002.

[24] Hernández, M. A., Rubio, M. J., y Ezquerro, J. A., *Secant-like methods for solving nonlinear integral equations of the Hammerstein type*, J. Comput. Appl. Math., **115**: 245–254, 2000.

[25] Hu, S., Khavanin, M., y Zhuang, W., *Integral equations arising in the kinetic theory of gases*, Appl. Anal., **34**: 261–266, 1989.

[26] Kantorovich, L. V., *The majorant principle and Newton's method* (en ruso), Dokl. Akad. Nauk. SSSR, **76**: 17–20, 1951.

[27] Kantorovich, L. V., y Akilov, G. P., *Functional analysis*, Pergamon Press, New York, 1982.

[28] Kincaid, D., y Cheney, W., *Análisis numérico: las mátemáticas del cálculo científicio*, Addison-Wesley Iberoamericana, Wilmington, 1994.

[29] Kreyszig, E., *Introductory functional analysis with applications*, John Wiley & Sons, New York, 1978.

[30] Laarsonen, P., *Ein abergradratisch konvergenter intertiver algorithmus*, Ann. Acad. Sci. fenn. Ser. A I, **405**: 1–10, 1969.

[31] Lindenstruss, J., y Tzafriri, L., *Classical Banach spaces I*, Springer Verlag, 1977.

[32] Lindenstruss, J., y Tzafriri, L., *Classical Banach spaces II*, Springer Verlag, 1979.

[33] Lund, J., y Vogel, C., *A Fully-Galerkin method for the solution of an inverse problem in a parabolic partial differential equation, numerical solution of an inverse*, Inverse Problems, **6**: 205–217, 1990.

[34] ORTEGA, J. M., y RHEINBOLDT, W. C., *Iterative solution of nonlinear equations in several variables*, Academic Press, 1970.

[35] OSTROWSKI, A. M., *Solution of equations and systems of equations*, Academic Press, New York, 1966.

[36] POTRA, F. A., *On a modified secant method*, Anal. Number. Theor. Approx., **8,2**: 203–214, 1979.

[37] POTRA, F. A., y PTÁK, V., *Nondiscrete induction and iterative processes*, Pitman, New York, 1984.

[38] RALL, L. B., *Computational solution of nonlinear operator equations*, Robert E. Krieger Publishing Company, Michigan, 1979.

[39] RASHIDINIA, J., y ZAREBNIA, M., *New approach for numerical solution of Hammerstein integral equations*, Appl. Math. Comput., **185**: 147–154, 2007.

[40] RUDIN, W., *Functional Analysis*, Alhambra, Madrid, 1979.

[41] SCHMIDT, J. W., *Eine übertragung der regula falsi auf gleichungen in Banachraum. I, II*. Z. Angew. Math. Mech., **43,1-8**: 97–100, 1963.

[42] SCHRÖDER, J., *Nichtlineare majoranten beim verfahrem der schrittweissen näherung*. Arch. Math. (Basel), **7**: 471–484, 1956.

[43] SERGEEV, A., *On the method of chords*. Sibirsk. Mat. Z., **2**: 282–289, 1961.

[44] STOKER, J. J., *Nonlinear vibrations*, Interscience-Wiley, New York, 1950.

[45] TRAUB, J. F., *Iterative methods for the solution of equations*, Prentice-Hall, Englewood Cliffs, New Jersey, 1964.

[46] ULM, S., *On the generalized divided differences. I, II*, Izv. Akad. Nauk Eston. SSR, Ser. Fitz. Mat., **16**: 146–156, 1967.

[47] VARONA, J. L., *Graphic and numerical comparison between iterative methods*, Math. Intelligencer, **24**: 37–46, 2002.

[48] WOJTASZCZYK, P., *Banach spaces for analysis*, Cambridge University Press, 1991.

[49] WOLFRAM, S., *The Mathematica book, 5th ed.*, Wolfram Media / Cambridge University Press, 2003.



Servicio de Publicaciones
Biblioteca Universitaria
C/ Piscinas, s/n
26006 Logroño (La Rioja)
Teléfono: 941 299 187

http://publicaciones.unirioja.es
www.unirioja.es